



UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

Essays on Political Economy and Industrial Organization

Autor:
Galina Zudenkova

Director:
Luis C. Corchón

Codirector:
Ignacio Ortuño Ortín

DEPARTAMENTO DE ECONOMÍA

Getafe, Junio de 2010

TESIS DOCTORAL

Essays on Political Economy and Industrial Organization

Autor: Galina Zudenkova

Director: Luis C. Corchón

Codirector: Ignacio Ortuño Ortín

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Vocal:

Vocal:

Secretario:

Calificación:

Getafe, de Junio de 2010

TABLE OF CONTENTS

LIST OF TABLES	iv
LIST OF FIGURES	v
ACKNOWLEDGEMENTS	vi
SUMMARY	viii
I SPLIT-TICKET VOTING: AN IMPLICIT INCENTIVE APPROACH	1
1.1 Introduction	1
1.2 Ticket Splitting in a Large City	6
1.2.1 Politicians	9
1.2.2 Voters	10
1.2.3 Equilibrium Concept	12
1.2.4 Equilibrium	13
1.2.5 Dynamics	17
1.3 Ticket Splitting in a Region with Small Municipalities	20
1.4 Empirical Model	24
1.4.1 Data Description	25
1.4.2 Empirical Results	27
1.5 Conclusion	28
1.6 Appendix	28
II SINCERE LOBBYING FORMATION	34
2.1 Introduction	34
2.2 A Simple Model of Public Goods Provision	39
2.3 Lobby Groups	40
2.4 Sincere Lobbying Formation	43
2.5 Numerical Solution: Lognormal Distribution of Income	47
2.6 Conclusion	51
2.7 Appendix	52

III COMPUTING WELFARE LOSSES FROM DATA UNDER IMPER-	
 FECT COMPETITION WITH HETEROGENEOUS GOODS	56
3.1 Introduction	56
3.2 The Linear Model	60
3.2.1 The Symmetric Case	60
3.2.2 Heterogeneous Firms	65
3.3 A Model of a Large Group	70
3.4 Conclusion	72
3.5 Appendix	73
IV WELFARE LOSSES IN MODELS OF HORIZONTAL AND VERTICAL	
 DIFFERENTIATION	78
4.1 Introduction	78
4.2 Horizontal Differentiation: The Hotelling Model	80
4.3 Horizontal Differentiation: The Salop Model	89
4.4 Vertical Differentiation	94
4.5 Conclusion	99
4.6 Appendix	100
REFERENCES	103

LIST OF TABLES

1	Ticket splitting (TS) and municipality population share.	27
2	Summary statistics of the predicted probability of ticket splitting $\hat{P}(y_{it} = 1)$	27
3	Sample characteristics and descriptive statistics.	33
4	The US gross income descriptives from LIS household files (year 2000), normalized descriptives, corresponding parameters for lognormal distribution, socially optimal tax rate and median voter preferred tax rate.	53
5	Numerical results and relative magnitudes of lobbying formation for utilitarian and pro-median governments with $\alpha = 1$, $\alpha = 100$ and $\alpha = 0$	53

LIST OF FIGURES

1	Mayor M and governor G reelection outcomes under linear rules (β_M, b_M) and (β_G, b_G) in the two-dimensional space of performances p_M and p_G	11
2	Best response functions of politicians i and j for independent reelections (black), positively correlated reelections (red) and negatively correlated reelections (blue) in states S and D	14
3	Politicians' reelection outcomes under equilibrium rules β_M^* and β_G^* in states S and D	19
4	Percentage of split-ticket voting (ST) in the regions included in the sample.	33
5	Sincere Lobbying Formation for US data: the case of utilitarian government with $\alpha = 1$	48
6	Sincere Lobbying Formation for US data: the case of pro-median government with $\alpha = 1$	49
7	Sincere indifferent poor π and rich ρ as a function of α for US data: π_U, ρ_U (utilitarian government) and π_M, ρ_M (pro-median government).	54
8	Equilibrium tax rate as a function of α for US data: t_U^* (utilitarian government) and t_M^* (pro-median government).	54
9	The size of the lobby of the poor s^P and the size of the lobby of the rich s^R as a function of standard deviation \sqrt{Var} for constant mean $\hat{x} = 5.611469$ with $\alpha = 1$: s_U^P, s_U^R (utilitarian government) and s_M^P, s_M^R (pro-median government).	55
10	Equilibrium tax rate as a function of standard deviation \sqrt{Var} for constant mean $\hat{x} = 5.611469$ with $\alpha = 1$: t_U^* (utilitarian government) and t_M^* (pro-median government).	55
11	PWL as a function of $\frac{\alpha-c}{\beta}$	86
12	PWL as a function of observable relative distance \mathfrak{r} from the market edges to brand locations when the market is not covered, $\mathfrak{m} < 1$ (dash), and when the market is covered, $\mathfrak{m} = 1$ (solid).	88
13	PWL as a function of $\frac{(\alpha-c)n^\gamma}{\beta}$ for linear $\gamma = 1$ (solid) and quadratic $\gamma = 2$ (dash) transportation costs.	92
14	PWL as a function of market coverage $\mathfrak{r}\mathfrak{n}$ for linear (solid) and quadratic (dash) transportation costs.	93
15	$\frac{s_1^*}{S}$ as a function of $\frac{b}{a} > 2$	97
16	PWL as a function of $\frac{b}{a} > 2$	98
17	PWL as a function of $\frac{p_1}{p_2}$ for $\mathfrak{m} = 1$ and $0 < \frac{p_1}{p_2} \leq \frac{2}{5}$	99
18	PWL as a function of \mathfrak{m} for $\frac{7}{8} < \mathfrak{m} < 1$ and $\frac{p_1}{p_2} = \frac{2}{7}$	100

ACKNOWLEDGEMENTS

I am deeply grateful to my advisor Luis C. Corchón and co-advisor Ignacio Ortuño Ortín for their guidance, encouragement and generous help. This thesis has significantly benefited from their insightful comments and discussions. In addition, Luis C. Corchón gave me a great opportunity to work with him on a joint research project, which led to the third and forth chapters of this thesis.

I am indebted to Anna Bogomolnaia, Antonio Cabrales, Marco Celentani, Jacques Crémer, M^a Ángeles de Frutos and Rebecca B. Morton for their detailed comments and support. They have carefully read and helped me to improve this thesis.

I received many helpful suggestions from faculty members and graduate students at the Department of Economics of the Universidad Carlos III de Madrid. In particular, I want to thank Pedro Albarrán, Vanessa Berenguer Rico, Raquel Carrasco, Klaus Desmet, Mikhail Drugov, Ramiro de Elejalde, Jesús Gonzalo, Albert Ma, Matilde Machado, Ricardo Mora, Diego Moreno, António Osório-Costa, Esther Ruiz, Antonio Romero-Medina and Carlos Velasco.

The research leading to this thesis received the financial support from the Spanish Ministry of Science and Innovation and was partially carried out while visiting the Department of Economics at New York University, to both of which I am grateful. I am especially indebted to my NYU faculty sponsor Raquel Fernández, whose comments on a preliminary version of the first chapter helped me to improve the interpretation of its main results.

The thesis has greatly benefited from useful comments and suggestions by Nikolay Archak, David Austen-Smith, Carmen Beviá, Steven Brams, Filipe Campante, Micael Castanheira, Alessandro Lizzeri, François Maniquet, Nicola Persico, Maria Petrova, François Salanié, James Snyder, Albert Solé, Pilar Sorribas and seminar and conference participants at Alabama, Barcelona, Bilbao, Bonn, Boston, Brussels, Cambridge, Evanston, Istanbul, London,

Louvain-la-Neuve, Madrid, Marseille, Milan, New York, Raleigh, Reus, Salamanca, San Sebastián, Siena, Silvaplana, Toulouse and Valencia.

Finally, I am most grateful to my family for their support and patience. I could not have done it without their help.

SUMMARY

This thesis comprises essays on Political Economy and Industrial Organization. The first two chapters are composed of research papers on Political Economy, which contribute to the literature on split-ticket voting and lobbying formation, respectively. The other two chapters on Industrial Organization are based on joint research papers with Luis C. Corchón, where we study the welfare losses yielded by imperfect competition under product differentiation. Each chapter can be considered independently of the rest.

Chapter I proposes a novel rationale for split-ticket voting, when citizens vote for candidates from different parties in simultaneous elections. It applies a political agency framework with implicit incentives to study ticket splitting in simultaneous municipal and regional elections. The results suggest that ticket splitting is a natural outcome of the optimal reelection scheme adopted by voters to motivate politicians' efforts in a retrospective voting environment. An office-motivated politician (mayor or governor) is assumed to prefer her counterpart to be affiliated with the same political party. This correlation of incentives leads the voters to adopt a joint performance evaluation rule, which is conditioned on the politicians belonging to the same party or different parties. The model is dynamic, generating predictions of split-ticket voting over time. Ticket splitting is shown to be less likely than electing candidates from the same party, but somewhat affected by ticket splitting in the previous period. Ticket splitting is also more likely in smaller municipalities, where the party affiliation of a mayor is assumed to be of less importance to the governor. These findings are consistent with empirical evidence from simultaneous municipal and regional elections held in Spain.

Chapter II analyzes the impact of lobbying, modeled as a common-agency problem, on a public goods provision. It introduces a sincere lobbying formation condition for equilibrium, namely, an equilibrium occurs only if no lobby member would prefer her lobby to cease to exist. The results suggest that individuals with more extreme income levels are more likely to join lobbying activities. The model is solved numerically for the US data to show that

lobbying does not necessarily favor the rich. If the government does not care about its reelection chances and does care about individuals' welfare, final policy outcome favors the poor. The lobby of the poor is more numerous and in total contributes more than the lobby of the rich. However, per member contribution is greater in the lobby of the rich. In the case where the elections are coming and the government wants to be reelected, lobbying does favor the rich. Although the lobby of the poor is more numerous, it contributes in total and per member less than the lobby of the rich. If the government cares only about contribution payments all individuals participate in lobbying, and political competition results in a socially optimal outcome.

Chapter III studies the percentage of welfare losses (PWL) yielded by imperfect competition under product differentiation. When demand is linear, even if prices, outputs, costs and the number of firms can be observed, PWL is arbitrary in both Cournot and Bertrand equilibria. If in addition the elasticity of demand (resp. cross elasticity of demand) is known, PWL in a Cournot (resp. Bertrand) equilibrium is calculated. When demand is isoelastic and there are many firms, PWL can be computed from prices, outputs, costs and the number of firms. The results suggest that price-marginal cost margins and demand elasticities may influence PWL in a counterintuitive way. The chapter also provides conditions under which PWL increases or decreases with concentration.

Chapter IV studies PWL in models of horizontal and vertical differentiation. In the Hotelling model, PWL is shown to depend on the underlying parameters in a non-monotonic way. It is also shown that PWL can be calculated from market data—locations and market size—except when the market is covered and exhibits maximal product differentiation. PWL can be very large—up to 37.4%—arising from firms located in the wrong places. In the Salop model, PWL can be calculated from market size. PWL may be large—up to 25%—but, in general, smaller than in the Hotelling model because firms are optimally located here. Finally, under vertical differentiation with two firms, PWL is discontinuous, but can be calculated from market prices and market coverage. In this model PWL is modest, always below 8.33%.

Esta tesis comprende cuatro ensayos de economía política y organización industrial. Los dos primeros capítulos se refieren a temas de economía política: Al problema del voto dividido y al de la formación de grupos de presión. Los otros dos capítulos tratan de temas de organización industrial y están basados en trabajos conjuntos con Luis C. Corchón. En ellos se estudian las pérdidas de bienestar generadas por la competencia imperfecta cuando el producto está diferenciado. Cada capítulo puede leerse independientemente del resto.

El Capítulo I propone una nueva explicación del voto dividido, cuando los ciudadanos votan en elecciones simultáneas por candidatos que podrían pertenecer a partidos políticos diferentes. Se usa un modelo de principal-agente con incentivos implícitos al voto dividido en elecciones municipales y regionales. Los resultados sugieren que el voto dividido es el resultado de un esquema de reelección óptimo que es adoptado por los votantes para motivar el esfuerzo de los políticos. Un político que pretende alcanzar la victoria electoral preferiría estar rodeado de políticos de su mismo partido. Esta correlación de incentivos lleva a los votantes a adoptar una regla de evaluación conjunta que está influida por el hecho de que los políticos pertenecen al mismo o a diferente partido. El modelo es dinámico, generando predicciones de voto dividido en el tiempo. Se prueba que este fenómeno es menos probable que elegir candidatos del mismo partido y está afectado por la división de votos en el período previo. La división del voto es más probable que ocurra en municipalidades pequeñas donde la afiliación política de un alcalde es de menos importancia al gobernador de la provincia. La evidencia empírica de elecciones municipales y regionales en España confirma esas predicciones teóricas.

El Capítulo II analiza el impacto de los grupos de presión en la provisión de bienes públicos cuando el problema se modeliza como uno de agencia común. Se define la noción de "Formación sincera de grupos de presión" en la que ningún miembro del grupo prefiere que su grupo no exista. Los resultados sugieren que los individuos con niveles de renta más extremos son los que tienen más incentivos para ser miembros de los grupos de presión. El modelo se resuelve numéricamente usando datos de Estados Unidos y muestra que los grupos de presión no necesariamente favorecen a los ricos. Si al gobierno en el poder no le

importa su reelección y le importa el bienestar de los ciudadanos, el resultado de la acción de los grupos de presión favorece a los pobres. El grupo de presión de estos últimos es más numeroso y contribuye con más recursos que el grupo de presión de los ricos. Sin embargo el grupo de presión de estos últimos tiene una mayor contribución por individuo. Si el gobierno desea ser reelegido, la actividad de los grupos de presión favorece a los ricos ya que estos contribuyen más en términos absolutos. Si al gobierno solo le importa el dinero que recibe de los grupos de presión el resultado es el óptimo socialmente.

El Capítulo III estudia el porcentaje de pérdida de bienestar que está generado por la competencia imperfecta cuando hay diferenciación del producto. Cuando la demanda es lineal el conocimiento de los precios la producción, los costes y el número de empresas no basta para obtener las pérdidas de bienestar y éstas son arbitrarias bajo competencia en cantidades o en precios. Pero si la elasticidad de la demanda (propia o cruzada) se conoce, entonces las pérdidas de bienestar se pueden deducir de los datos de mercado cuando la competencia es en cantidades o en precios. Cuando la demanda es isoelástica, las pérdidas de bienestar se pueden computar de los precios, la producción, los costes y el número de empresas. En algunos casos, los márgenes precio-coste marginal y las elasticidades de demanda pueden afectar a las pérdidas de bienestar de manera muy poco intuitiva. También se estudian las condiciones bajo las que las pérdidas de bienestar aumentan o disminuyen con la concentración.

El Capítulo IV estudia las pérdidas de bienestar en modelos de diferenciación horizontal y vertical. En el caso del modelo de Hotelling, las pérdidas de bienestar dependen de los parámetros que definen la economía de una manera que no es monótona. También se prueba que las pérdidas de bienestar se pueden calcular de los datos de mercado como la localización y el tamaño del mercado, excepto cuando todo el mercado está cubierto y hay diferenciación máxima. Las pérdidas de bienestar pueden ser muy grandes, de hasta el 37.4% y son debidas principalmente a la localización incorrecta de las empresas. En el modelo de Salop, las pérdidas de bienestar se pueden calcular del tamaño de mercado. Aquellas pueden ser grandes, de hasta un 25%, pero son, en general, más pequeñas que en el modelo

de Hotelling, ya que en este caso las empresas están situadas óptimamente. Finalmente bajo diferenciación vertical con dos empresas, las pérdidas de bienestar son discontinuas pero pueden ser calculadas de los precios y la cobertura del mercado. En este modelo, las pérdidas de bienestar son modestas, siempre más pequeñas que el 8.33%.

CHAPTER I

SPLIT-TICKET VOTING: AN IMPLICIT INCENTIVE APPROACH

1.1 Introduction

Split-ticket voting, when a citizen votes for candidates from different parties in simultaneous elections, is a common feature of modern political systems. Ticket splitting has mostly been studied in the context of the US, where simultaneous presidential and congressional elections are held every four years (see Burden and Kimball 2002, Fiorina 1996, Jacobson 1990, and Zupan 1991, among many others).

To the best of my knowledge, there are several formal models of split-ticket voting. They are mainly based on an institutional assumption that a final policy choice depends on both the executive and the composition of the legislature. Indeed, Alesina and Rosenthal (1995, 1996) elaborated on the policy balancing argument, showing that citizens strategically split tickets to avoid the extreme policies that may arise when the executive and legislative branches are allied. Chari et al. (1997) built a model on the budgetary externality of concentrated government spending financed by uniform taxes. They found that voters prefer a fiscally conservative president (to restrain overall spending) and fiscally liberal congressmen (who promote spending in their home districts). In turn, Bugarin (2003) showed that voters split tickets in order to reinforce opposition in the legislature as a means of the executive corruption control. Fox and Van Weelden (2009) proposed an alternative rationale for ticket splitting based on the need for effective oversight of the executive in a career concerns framework. Note that these models explain ticket splitting in a particular institutional setting where a final policy outcome is determined by both branches of government (e.g., the executive and the legislature). This paper complements the aforementioned

literature and analyzes ticket splitting at lower levels of government—in particular, municipal and regional levels of government—where the assumption about such an institutional setting can be relaxed. As a rule, mayors and governors face distinct well-determined tasks and have to implement distinct policies.

In this paper, I apply an implicit incentive approach to explain ticket splitting within a retrospective voting model (i.e., a political agency model with moral hazard). In my framework, split-ticket voting arises as an outcome of the optimal implicit reward scheme voters use to induce politicians' efforts. The model is dynamic, generating predictions of ticket splitting over time. This feature is original; none of the aforementioned contributions has analyzed ticket splitting in a dynamic context.

I consider a representative municipality in a region where the mayoral and gubernatorial elections are simultaneous. I work with the two settings: one with a single large city whose vote is decisive for the outcome of the regional election, and one with many cities of varying size. In the latter case, each city has a probability proportional to its population of playing a pivotal role in the regional election.

I use a political agency model of interaction between politicians and their constituency in the presence of a moral hazard problem. The politicians want to be reelected for another term, and are held accountable for their past performance at the moment of election. The politicians therefore have incentives to satisfy the voters' wishes. In addition, I assume that the politicians are loyal to their respective political parties: the mayor prefers the governor to be affiliated with the same political party, and vice versa.¹ Hence, the incentives of the mayor and governor are correlated. The voters care about the politicians' performances, which are observable but not contractible. The voters evaluate the incumbents' performance and vote accordingly. More precisely, the voters employ implicit evaluation rules when deciding whether to reward (reelect) politicians. Obviously, voters can influence the politicians' behavior through their choice of evaluation rules. I restrict the space of possible

¹Fox and Van Weelden (2009) introduce a similar assumption about the partisan preferences of the legislature. In particular, in their career concerns setup the legislature ("overseer") can care about the executive's reputation. For example, a partisan overseer may seek to damage the reputation of an executive from the other party while seeking to protect the reputation of an executive from her own party.

evaluation rules to linear functions of performance. The evaluation rules are also required to be sequentially rational.

I show that given the correlation between the two politicians' incentives, voters are better off adopting a *joint* performance evaluation rule (conditioned on the incumbents belonging to the same party or different parties) rather than an *individual politician* performance evaluation rule. In particular, the voters evaluate the performance of the mayor and governor from the same party as a team. If the mayor and governor belong to different parties, then the voters compare their performances to create a competitive environment. This combination of rules implies that improved performance increases a politician's own reelection probability, while increasing/decreasing the reelection probability of her partisan ally/rival in the other office. Politicians therefore have extra incentives to perform better, for the sake of their party as well as for themselves.

In equilibrium, the reelection outcomes of incumbents from the same party are therefore positively correlated: voters tend to reward both incumbents from a well-performing party, or punish both incumbents from a poorly-performing party. The reelection outcomes of incumbents from different parties are negatively correlated: voters tend to reward the incumbent from the better-performing party while punishing the other incumbent. This combination generates a dynamic of partisan voting: whether or not the incumbent politicians belong to the same party, ticket splitting is always less likely than electing both candidates from the same party.

Next, I consider two cases. First, there might be some preference for incumbents such that a mayor/governor prefers a known incumbent from her own party to a new politician affiliated with the same party to hold the other office. The allied incumbents therefore have extra incentives to perform well so that they continue working together. This implies that politicians from the same party exert a higher total effort than politicians from different parties. Voters adopt joint performance evaluation rules, under which the reelection outcomes of incumbents from the same party are somewhat more correlated in absolute value, as compared with the ones from different parties. Ticket splitting is therefore more likely in elections where the incumbents belong to different parties, which in turn is a consequence

of ticket splitting in the previous period. Second, there might be some preference for newcomers such that a mayor/governor prefers a new unknown member of her own party to a known incumbent ally for the other office. So the incumbents from the same party have somewhat less incentives to perform well and therefore exert a lower total effort than politicians from different parties. The joint reelection rules for politicians from the same party are less correlated (in absolute value) than the ones for politicians from different parties. Ticket splitting is thus more likely in elections where the incumbents belong to the same party, i.e., where the voters did not split tickets in the previous period.

These results rest on the assumption of politicians' *party loyalty*; that is, I assumed that a mayor/governor cares about the party affiliation of the governor/mayor. The joint performance evaluation rules then give extra implicit incentives to the politicians. If I relax the assumption of party loyalty, this effect vanishes and the voters no longer evaluate incumbents jointly. Instead they use a cut-off rule that each incumbent is reappointed only when her individual performance exceeds a critical threshold. In this situation, the probability of ticket splitting goes up.

Furthermore, I assume that the mayor's party affiliation is of less importance in smaller municipalities. This may happen for two reasons: the mayors are more likely to run as independents (or be affiliated with a minor party) and the constituencies already know the candidates very well. (Recall that one of the major roles of political parties is to provide information about unknown politicians.) Thus, a governor cares less about the party affiliation of small-town mayors. As a result, the two politicians' incentives are less correlated. In this situation, the incumbents are more likely to be evaluated individually rather than jointly, which increases the probability of ticket splitting.

To sum up the predicted dynamics of the model, I find that ticket splitting is less likely than voting for candidates from the same party. Moreover, ticket splitting is somewhat affected by ticket splitting in the previous period. Finally, I find that ticket splitting is more likely to occur in small municipalities than in large ones. In the empirical section of this paper, I estimate the probability of ticket splitting using panel data on the aggregate results of simultaneous municipal and regional elections in Spain. The predictions outlined

above are consistent with the results of this empirical analysis.

I turn now to the fundamental question of why political constitution is modeled as political agency. Firstly, in addition to a sound theoretical framework, this approach has received considerable empirical support (see, for example, Peltzman 1992 and Besley and Case 1995a, 1995b, 2003). Besley (2006) provides an excellent introduction to political agency models and "emphasizes the empirical potential of these models in explaining real world policy choices."² Secondly, I believe that at the municipal and regional levels politicians' tasks require mainly managerial skills. This view is supported by the empirical work of Ferreira and Gyourko (2009), who found that in US cities the mayor's party affiliation does not affect the size of the city government and the allocation of spending. In a recent article in the *New York Times*, Glaeser points out that "lack of ideology has become a major feature of big city mayors... They are... managerial mayors, appreciated by voters because they succeed in making the city work."³ The political agency approach may therefore be appropriate to model local political constitutions. Even so, elected politicians can only be offered implicit incentive schemes; public policies are difficult to reward with explicit contracts.

The retrospective voting model I use goes back to Barro (1973). Ferejohn (1986) extended the model and studied subgame-perfect equilibria rather than Nash equilibria. Persson et al. (1997) use a retrospective voting approach to study rent extraction. In Austen-Smith and Banks (1989) voters adopt retrospective voting strategies that are conditioned on the difference between the incumbent's performance and her initial policy platform. Banks and Sundaram (1993, 1996) analyze retrospective voting settings with both moral hazard and adverse selection, and with term limits respectively.

The results of this paper are also related to the literature on horizontal and vertical intergovernmental competition. Most analyses of horizontal competition are based on the assumption of interjurisdictional mobility of consumers, à la Tiebout (1956). In a similar

²Besley (2006), p. 3.

³Edward L. Glaeser "Lower (and More Realistic) Presidential Expectations," January 20, 2009. Available online at <http://economix.blogs.nytimes.com/2009/01/20/lower-and-more-realistic-presidential-expectations/> (accessed January 26, 2009).

vein, the literature on yardstick competition between jurisdictions started with the seminal work of Salmon (1987), to be followed by Besley and Case (1995a), Bordonon et al. (2004), Sand-Zantman (2004), Belleflamme and Hindriks (2005), Besley and Smart (2007) and others. The main assumption of this literature is that under decentralization, voters use a comparative performance evaluation between different local governments to create yardstick competition.

The vertical competition literature, on the other hand, assumes that "senior and junior governments provide similar or comparable services, and that office-holders in the government which is judged by citizens to be the more efficient supplier will increase their probability of getting the vote of these citizens"⁴ (see Breton 1996, Breton and Fraschini 2003, Breton and Salmon 2001, Volden 2005 and Volden 2007). I follow these authors in assuming that voters compare the performance of local and regional governments, and are likely to reward the more efficient politicians with reelection. There is, however, an important difference between my research and the papers just cited. In the intergovernmental competition literature, the comparative performance evaluation result is driven by either correlated shocks or interjurisdictional spillover. In my model, the joint performance evaluation arises from the fact that the politicians' incentives are correlated: each one cares not only about her own reelection prospects, but also about the success of other politicians affiliated with the same political party.

The remainder of the paper is organized as follows. Section 1.2 lays out a model of ticket splitting in one large city pivotal to the regional election. Section 1.3 presents a model of ticket splitting in a region with many small cities. Section 1.4 describes empirical results on Spanish elections. Finally, Section 1.5 concludes.

1.2 Ticket Splitting in a Large City

In this section I study ticket splitting in simultaneous municipal and regional elections held in a large city. I assume that the city is large enough that its vote will be decisive for the regional election.

⁴Breton and Salmon (2001), p. 139.

Consider a large city, with an infinite horizon, that has to elect mayor M (for the municipal government) and governor G (for the region to which this city belongs). The city is inhabited by a large number (formally a continuum) of individuals. The individuals live forever. At the beginning of each period, the elections take place simultaneously and the winners are determined by majority rule. Politicians running for both offices belong to one of the two political parties, L or R . I assume that there is exactly one candidate from each party—the incumbent and an opponent—in each election and in each period. The opponents are identical to the incumbents in all respects except party affiliation. The participation constraints of the politicians are always satisfied, and there is no term limit.

First I will describe a stage game with a stationary environment, where time subscripts can be dropped with no risk of confusion. I will then consider an infinitely repeated game.

One stage is a sequential political agency game between politicians (the mayor and governor) and their constituency (the voters). While in office, each politician $i \in \{M, G\}$ has to implement a policy determined by her unobservable effort a_i .⁵ The set of efforts available to each politician is taken to be a non-degenerate interval $[0, \bar{a}] \subset \mathbb{R}$. I assume that the performance of politician i , p_i , is observed with an independent and unobservable noise ε_i :

$$p_i = a_i + \varepsilon_i,$$

with $\varepsilon_i \sim N(0, \sigma^2)$.⁶

The reward of politician i is denoted by $\Pi_i(a_i)$. Effort is costly, and I assume the standard convex cost function $\frac{a_i^2}{2}$.⁷ The mayor and governor independently choose effort

⁵One can add an adverse selection problem by assuming that policy outcomes are determined by effort and ability. The results are qualitatively the same if ability is modeled as a moving average (to capture the idea that a policymaker's competence changes slowly over time).

⁶I have an extended version of the model, available upon request, where the two noise terms ε_M and ε_G are correlated and follow a bivariate normal distribution. I want to concentrate however on the case where voters introduce joint performance evaluation due to the correlation between politicians' incentives rather than the correlation between shocks. The latter topic has been widely studied in the context of team evaluation in contract theory (for an overview, see Bolton and Dewatripont 2005) and in the literature on yardstick competition (see the references on yardstick competition in the Introduction).

⁷I have an extended version of the model, available upon request, where the cost of policy implementation for the mayor and governor from the same party is different than for the politicians from rival parties (e.g., because of synergy). The results of this extended model are qualitatively the same.

levels a_i to maximize their utility, which is given by

$$\Pi_i(a_i) - \frac{a_i^2}{2}.$$

The function $\Pi_i(a_i)$ will be explicitly defined in subsection 1.2.1.

There is no cost of voting, and I assume that there are no abstainers. The individuals differ in their preferences over political parties. I assume that some individuals always prefer party L to party R , and vote for candidates from party L in both elections. Likewise, other individuals are loyal to party R . However, there is a large group of uncommitted individuals, whose votes are decisive for the outcome of both elections. These voters care about the politicians' performances in each period according to a linear utility function

$$p_M + p_G.$$

In what follows, I refer to this group of decisive voters simply as "the voters".

I assume that the voters coordinate to apply the same retrospective reappointment rules to reelect the incumbents. I follow the literature (e.g. Persson et al. 1997) in restricting the strategy space such that the voters base their reappointment decision solely on the politicians' performances in the current period, not in any previous period. See Persson et al. (1997) for a discussion of the plausibility of this approach, and Fair (1978) and Kramer (1971) for empirical findings in its favor.

Denote by S the state where mayor M and governor G are members of the same party (either L or R), and by D the state where M and G belong to different parties. State S and state D occur when in the previous period the voters did not split or split tickets respectively.

The timing of events in the stage game is as follows. First, state S or D is realized. Second, the voters decide on the reappointment rules to be used in the coming elections. Third, the politicians exert efforts a_M and a_G . Fourth, the politicians' performances p_M and p_G are observed. Finally, both elections take place simultaneously and the voters apply the selected reappointment rules to reward or punish the incumbents.

In the following subsection I describe the politicians' preferences. I will then turn to the voters' problem and define an equilibrium concept.

1.2.1 Politicians

The politicians' preferences are as follows. First, once elected, mayor M and governor G want to be reelected in the next period. Moreover, M wants to improve her party's chances to win the governor's office in the coming election. If G and M belong to the same party, then M prefers G to be reelected. Otherwise, M wants the voters to appoint a new governor (from her own party) for the next term. Likewise, G wants M to be reelected if they are members of the same party, and wants the opponent to be appointed if M is from the rival party. The value of holding office is normalized to 1. The values, which M and G associate to their parties' holding the other office, are denoted by λ_M and λ_G respectively. Furthermore, denote by $\Pr_i(\cdot)$ the probability of being reelected to office $i \in \{M, G\}$ in the coming election. Therefore, politician i has the following reward function $\Pi_i : [0, \bar{a}]^2 \rightarrow \mathbb{R}$ that depends continuously on both politicians' efforts:

$$\Pi_i(a_i, a_j) = \begin{cases} \Pr_i(a_i, a_j) + \lambda_i^S \Pr_j(a_i, a_j) & \text{if } S \\ \Pr_i(a_i, a_j) + \lambda_i^D (1 - \Pr_j(a_i, a_j)) & \text{if } D, \end{cases}$$

where $i, j \in \{M, G\}$ and $j \neq i$. The preferences stated above reflect the politicians' allegiance to their respective parties; individual politicians care about their party's overall representation in mayor and governor offices, not just their own reelection prospects.⁸ Still, the reasonable assumption here is that a mayor/governor values her own office more than her party's representation in the other office; i.e., $0 \leq \lambda_i \leq 1$.⁹ I call λ_i the degree of politician i 's loyalty to her party (or the strength of her party alignment). The two alternative cases will be considered. First, there might be some preference for incumbents: $\lambda_i^S \geq \lambda_i^D$. This case reflects the idea that a mayor/governor might prefer an incumbent affiliated with the same party to an unknown candidate for the other office. Second, politicians might prefer newcomers: $\lambda_i^S < \lambda_i^D$. Thus, a mayor/governor would like a new politician (i.e., a newcomer) from the same party to be elected for the other office.

⁸ Alternatively, the stated preferences could arise because the mayor and governor have to interact while in office. Each prefers working with a member of her own party rather than a rival.

⁹ In other words, politician i does not mind reducing her reelection chances by 1% in exchange for increasing her ally's election probability by $\frac{1}{\lambda_i}\% \geq 1\%$.

The model can be generalized to the case where the incumbents maximize their intertemporal utility function (as in Ferejohn 1986). At this stage, however, I want to concentrate on the interactions between politicians and voters. I therefore assume, as in Alesina and Tabellini (2008), that the incumbents are myopic: they care about reelection only in the next period, not in any subsequent period.

1.2.2 Voters

The politicians' performances p_M and p_G (but not their composition between effort and noise) are observed at the end of each period but are not contractible. Public policies are difficult to reward with explicit contracts. It is more natural to use implicit incentive contracting in this situation. I assume that the voters coordinate on the same retrospective voting rule, and that there is no coordination failure among the voters. A coordination problem is a serious issue, but lies beyond the scope of the paper.

The voters observe politicians' performances p_M and p_G , and in the coming elections they reward incumbents according to their performances in the current period; i.e., they reappoint incumbents who have shown "good" results in the current period. A politician thrown out of office is never reappointed. In this case an opponent from the rival party is elected.

Obviously the voters can influence the politicians' behavior through their choice of evaluation rules. Intuitively, since politicians care about the reelection chances of each other, the reward rules should allow for joint performance evaluation. Under joint performance evaluation the voters condition reelection of politician i on her own performance p_i (to give her incentives to perform well since she wants to be reelected) and on performance of politician j , p_j (to give incentives to politician j since he cares about i 's reelection chances). I restrict the functional space of performance evaluation rules to linear joint evaluation rules (β_M, b_M) and (β_G, b_G) . β_M and β_G are the slopes of the mayor and governor evaluation rules respectively, while b_M and b_G are the corresponding intercepts; $\beta_M, \beta_G, b_M, b_G \in \mathbb{R}$, $\beta_M \beta_G \leq 1$. Under rules (β_i, b_i) , $i \in \{M, G\}$, the probability of being reelected for office i is

$$Pr_i(a_i, a_j) = P(\{p_i(a_i) + \beta_i p_j(a_j) \geq b_i\})$$

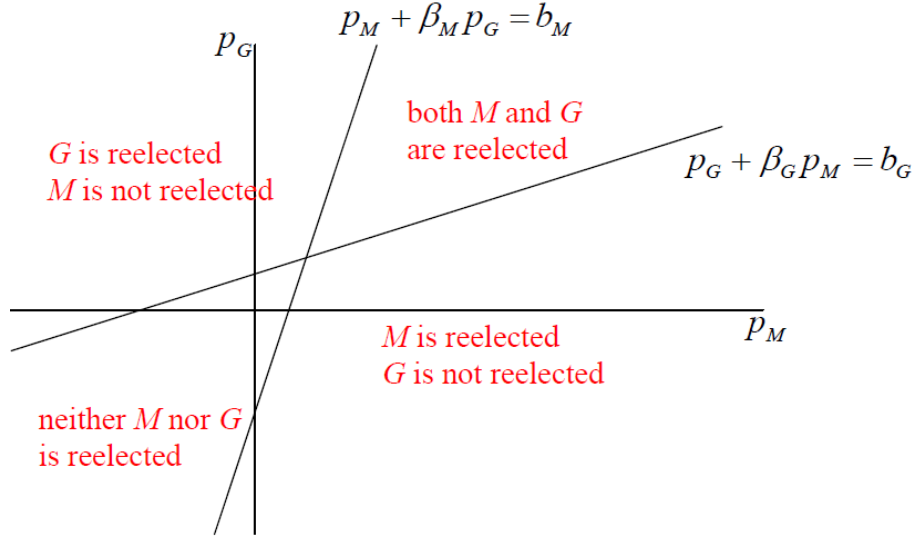


Figure 1: Mayor M and governor G reelection outcomes under linear rules (β_M, b_M) and (β_G, b_G) in the two-dimensional space of performances p_M and p_G .

with $i, j \in \{M, G\}$ and $j \neq i$. Figure 1 depicts the possible outcomes for M and G under rules (β_M, b_M) and (β_G, b_G) in the two-dimensional space of observed performances p_M and p_G . Note that I require $\beta_M \beta_G \leq 1$, so that line $p_M + \beta_M p_G = b_M$ is steeper than line $p_G + \beta_G p_M = b_G$. Otherwise, as one can see from Figure 1, a mayor and governor with poor performance would be reelected while politicians with better performance would not.

Note that under linear rules (β_M, b_M) and (β_G, b_G) , M 's reelection is determined by random variable $\varepsilon_M + \beta_M \varepsilon_G \sim N(0, (1 + \beta_M^2) \sigma^2)$, and G 's reelection is determined by random variable $\varepsilon_G + \beta_G \varepsilon_M \sim N(0, (1 + \beta_G^2) \sigma^2)$. I say that the two reelection events are independent when $\beta_M = 0$ and $\beta_G = 0$, positively correlated when $\beta_M > 0$ and $\beta_G > 0$, and negatively correlated when $\beta_M < 0$ and $\beta_G < 0$.

Since differentiable functions are linear in first-order approximation, the restriction to linear rules gives an approximate fit to more general evaluation rules. Furthermore, the linear evaluation rules allow for a closed-form solution. Linear approximation methods are widely used in macroeconomics to search for time-consistent equilibria (e.g. Krusell et al. 1997). In the contract theory literature, linear contracts have been shown to be optimal under some realistic assumptions (for an overview, see Bolton and Dewatripont 2005).

As I mentioned above, in this framework split-ticket voting emerges naturally from the

chosen evaluation rules. That is, following the chosen rules can result in the election of politicians from different parties. Henceforth I will find it more convenient to refer to S and D as the states characterized by the politicians belonging to the same party or different parties, keeping in mind that state S or state D occurs when the voters did not split tickets or split tickets respectively.

1.2.3 Equilibrium Concept

In the stage game I search for a subgame perfect equilibrium by analyzing the game backwards. First, I solve for the politicians' efforts a_M and a_G under rules (β_M, b_M) and (β_G, b_G) . Second, I examine the voters' choice of evaluation rules (β_M, b_M) and (β_G, b_G) . In what follows I introduce two definitions.

Given linear performance evaluation rules (β_M, b_M) and (β_G, b_G) , the equilibrium in effort strategies is a profile of efforts (a_M^e, a_G^e) such that

$$\Pi_i(a_i^e, a_j^e) - \frac{a_i^{e2}}{2} \geq \Pi_i(a_i, a_j^e) - \frac{a_i^2}{2} \text{ for each } a_i \in [0, \bar{a}],$$

where $i, j \in \{M, G\}$, $i \neq j$.

The voters are rational, so they realize that the only alternative to reelecting incumbents is voting for opponents from rival parties. The politicians' performances are additively separable in effort and noise, and all politicians behave in the same way irrespective of the noise. If elected, opponent i will exert equilibrium effort a_i^e , which maximizes her expected utility. Thus, the voters compare the incumbents' performances with their opponents' expected performances and vote accordingly. Formally,

$$b_i = a_i^e + \beta_i a_j^e.$$

Thus, the linear joint performance evaluation rules are solely determined by slopes β_i and β_j .

I define an equilibrium in rule strategies as the doublet (β_M^*, β_G^*) such that

$$a_M^e(\beta_M^*, \beta_G^*) + a_G^e(\beta_M^*, \beta_G^*) = \max_{\substack{\beta_M, \beta_G \\ \beta_M \beta_G \leq 1}} a_M^e(\beta_M, \beta_G) + a_G^e(\beta_M, \beta_G),$$

where $(a_M^e(\cdot), a_G^e(\cdot))$ is an equilibrium in effort strategies. Finally, the politicians' equilibrium efforts are denoted by $a_i^* \equiv a_i^e(\beta_i^*, \beta_j^*)$, $i, j \in \{M, G\}$, $i \neq j$.

Now consider an infinitely repeated game where each stage is the sequential political agency game presented above. First, recall that the reappointment decision depends only on the politicians' performances in the current period and not in any previous period. Second, recall that the incumbents are "myopic": they care about reelection only in the next period and not in any subsequent period. Under these assumptions on the voters' strategy space and the politicians' preferences, I consider the particular Markov Perfect Equilibrium of the infinitely repeated game where a stage game equilibrium in rule strategies is replicated infinitely often. The payoff-relevant states are S and D .

I must stress here that the evaluation rules are required to be sequentially rational; no precommitment is allowed. The model parameters are common knowledge, so the politicians know whether the voters used the evaluation rules that had been rational for them in the previous period or deviated. In the latter case, the politicians conclude that the voters reappoint incumbents randomly or use unknown rules that are not based on performance. As a result, from that period onward, the politicians will exert zero effort to minimize their costs. The voters know this, so they have no incentives to deviate and always reward incumbents according to the chosen rules.

1.2.4 Equilibrium

First, consider one stage game. Let the voters use evaluation rules (β_i, b_i) such that $b_i = a_i^e + \beta_i a_j^e$. Under these rules the politician i 's utility is

$$\Pi_i(a_i, a_j) - \frac{a_i^2}{2} = \begin{cases} P(\{p_i(a_i) + \beta_i p_j(a_j) \geq b_i\}) + \lambda_i^S P(\{p_j(a_j) + \beta_j p_i(a_i) \geq b_j\}) - \frac{a_i^2}{2} & \text{if } S \\ P(\{p_i(a_i) + \beta_i p_j(a_j) \geq b_i\}) + \lambda_i^D (1 - P(\{p_j(a_j) + \beta_j p_i(a_i) \geq b_j\})) - \frac{a_i^2}{2} & \text{if } D. \end{cases}$$

Politician i chooses effort a_i before observing realization of the noise, and takes the voters' expectations as given. Figure 2 depicts the politicians' best response functions for

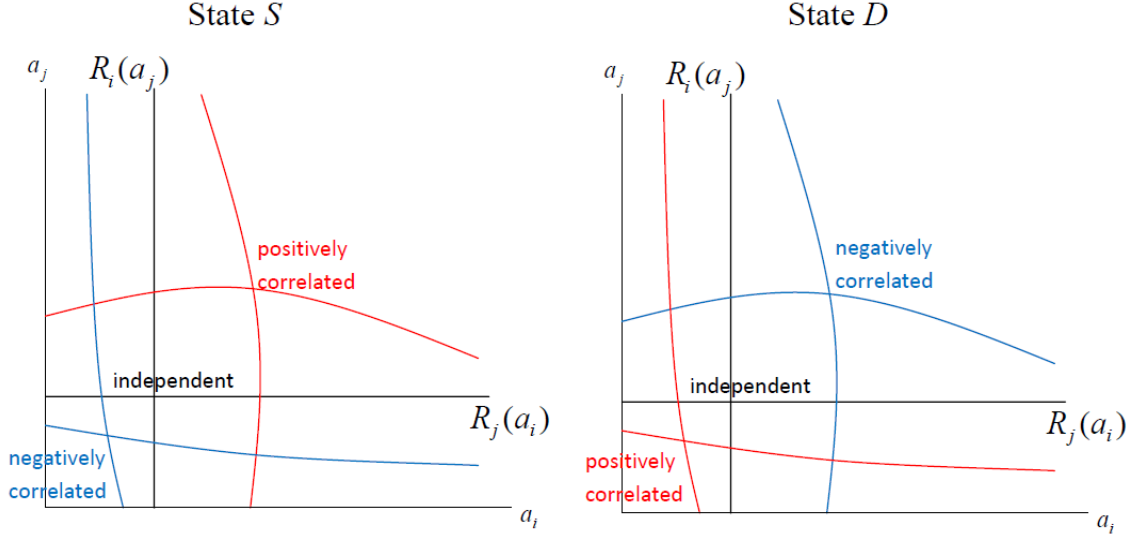


Figure 2: Best response functions of politicians i and j for independent reelections (black), positively correlated reelections (red) and negatively correlated reelections (blue) in states S and D .

independent reelections with $\beta_i = 0$ and $\beta_j = 0$ (in black), positively correlated reelections with $\beta_i > 0$ and $\beta_j > 0$ (in red) and negatively correlated reelections with $\beta_i < 0$ and $\beta_j < 0$ (in blue) for states S and D . Note that for independent reelections (in black) the best responses are flat in both states (since each politician's reelection depends only on her own effort). For positively correlated reelections (in red) the best responses shift upwards in state S and downwards in state D . Intuitively, under positively correlated reelections a politician has extra incentives to exert effort in state S (to increase her ally's reelection chances) and less incentives in state D (not to help her rival to get reelected). Finally, for negatively correlated reelections (in blue) the best responses shift downwards in state S and upwards in state D . Indeed, under negatively correlated reelections a politician does not want to damage her ally's reelection prospects and thus exert a lower effort in state S . She has however extra incentives to exert effort in state D to cut her rival's reelection chances. Note that there is a free-riding effect under positively correlated reelections in state S . Intuitively, politician i might prefer to exert a lower effort (and save the effort cost) if her partisan ally j performs well enough to help her to be reelected.

The result below establishes the existence of an equilibrium in effort strategies. Proofs

of this and other propositions are given in the Appendix.

Proposition 1 *Under linear performance evaluation rules β_M and β_G with $\beta_M\beta_G \leq 1$, there exists an equilibrium in effort strategies (a_M^e, a_G^e) given by*

$$a_i^e = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \left(\frac{1}{\sqrt{1+\beta_i^2}} + \frac{\lambda_i^S \beta_j}{\sqrt{1+\beta_j^2}} \right) & \text{if } S \\ \frac{1}{\sqrt{2\pi}\sigma} \left(\frac{1}{\sqrt{1+\beta_i^2}} - \frac{\lambda_i^D \beta_j}{\sqrt{1+\beta_j^2}} \right) & \text{if } D, \end{cases}$$

where $i, j \in \{M, G\}$, $i \neq j$.

Turn now to the voters' choice of evaluation rules β_M and β_G . Maximizing $a_M^e + a_G^e$ with respect to β_M and β_G yields an equilibrium in rule strategies (β_M^*, β_G^*) . I summarize the results in the following proposition (the proof is straightforward).

Proposition 2 *There exists an equilibrium in rule strategies (β_M^*, β_G^*) given by*

$$\beta_i^* = \begin{cases} \lambda_j^S & \text{if } S \\ -\lambda_j^D & \text{if } D, \end{cases} \quad (1)$$

where $i, j \in \{M, G\}$, $i \neq j$. The politicians' equilibrium efforts a_i^* are equal to

$$a_i^* = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \left(\frac{1}{\sqrt{1+(\lambda_j^S)^2}} + \frac{(\lambda_i^S)^2}{\sqrt{1+(\lambda_i^S)^2}} \right) & \text{if } S \\ \frac{1}{\sqrt{2\pi}\sigma} \left(\frac{1}{\sqrt{1+(\lambda_j^D)^2}} + \frac{(\lambda_i^D)^2}{\sqrt{1+(\lambda_i^D)^2}} \right) & \text{if } D. \end{cases} \quad (2)$$

According to Proposition 2, if politician j is loyal to her political party (i.e., $\lambda_j \neq 0$), the voters adopt a joint performance evaluation rule to reelect politician i . The probability of being reelected to office i under the joint rule is equal to

$$Pr_i(a_i, a_j) = \begin{cases} P\left(\left\{p_i(a_i) + \lambda_j^S p_j(a_j) \geq a_i^* + \lambda_j^S a_j^*\right\}\right) & \text{if } S \\ P\left(\left\{p_i(a_i) - \lambda_j^D p_j(a_j) \geq a_i^* - \lambda_j^D a_j^*\right\}\right) & \text{if } D. \end{cases}$$

Intuitively, the incentives of a mayor and governor are correlated, because they care about the overall representation of their party in both offices. The voters therefore reward politicians jointly rather than separately.

If the politicians belong to the same political party (state S), then the voters use a joint rule under which the reelection of politician i is positively correlated with the performance of politician j ($\beta_i > 0$). As a result, the voters evaluate the performance of the politicians from the same party as a team and tend to reward the incumbents from a well-performing party while punish the incumbents from a badly-performing party.

However, if the politicians belong to different parties (state D), the voters use a joint rule under which the reelection of politician i is negatively correlated with the performance of politician j ($\beta_i < 0$). As a result, the voters compare the performance of one politician to that of the other, creating a competitive environment between the parties. In this scenario the voters tend to reward the incumbent from the better-performing party, while punishing the incumbent from the worse-performing party.

In sum, due to the correlation between the mayor's and governor's incentives such that they care about their *party* chances of holding office, the voters are better off adopting *party* performance evaluation rather than *individual* performance evaluation.

Furthermore, the more loyal politician j is to her political party (the higher λ_j is), the more correlated the optimal reward scheme for politician i is with the performance of politician j (positively if S or negatively if D). Intuitively, if the politicians care equally about their own reelection chances and their party's election chances, then the best reward scheme would be perfectly correlated: in state S the incumbents are always reelected or dismissed together, while in state D reelection of one implies dismissal of the other.

The less loyal the politicians are to their political parties, the less correlated their incentives. As a result, the voters adopt the less correlated reelection rules in equilibrium. If politician j is not at all loyal to her political party ($\lambda_j = 0$), then the optimal rule to reappoint politician i is a simple cut-off rule: she is reappointed only if her observed performance exceeds a critical threshold given by the equilibrium effort for this office. That is, the probability of being reelected to office i depends only on a_i :

$$Pr_i(a_i) = P(\{p_i(a_i) \geq a_i^*\}).$$

Intuitively, when politicians care only about their own reelection prospects, the voters are

better off rewarding politician's individual performance rather than the party's performance.

Next, compare the equilibrium efforts of politicians from the same party and from different parties, as given in (2). If there is some preference for incumbents (i.e., $\lambda_i^S \geq \lambda_i^D$) then politicians from the same party exert a higher total effort than politicians from different parties:

$$(a_M^* + a_G^*)|_S \geq (a_M^* + a_G^*)|_D \text{ if } \lambda_M^S \geq \lambda_M^D \text{ and } \lambda_G^S \geq \lambda_G^D.$$

Intuitively, politicians from the same party are more loyal to their political parties than politicians from different parties, so the former have extra incentives to exert a higher effort (in order to increase the probability of their counterparts' reelection). If a mayor and governor prefer newcomers (i.e., $\lambda_i^S < \lambda_i^D$) then politicians from the same party have less incentives to perform well. So politicians from the same party exert a lower total effort than politicians from different parties:

$$(a_M^* + a_G^*)|_S < (a_M^* + a_G^*)|_D \text{ if } \lambda_M^S < \lambda_M^D \text{ and } \lambda_G^S < \lambda_G^D.$$

How do the equilibrium efforts a_i^* in (2) depend on parameters' values? First, larger variance of noise σ^2 decreases the politicians' efforts. Intuitively, more randomness in the observed performances p_M and p_G makes the reelection probabilities less sensitive to effort, reducing the politicians' incentives. Second, if politician i 's party alignment λ_i is strengthened, the equilibrium effort of politician i , a_i^* increases while that of politician j , a_j^* decreases. The more politician i cares about her ally's appointment to office j , the more incentives she has to perform better. However, this weakens politician j 's incentives to exert effort, because his reelection becomes less sensitive to his own effort.

In the infinitely repeated game, one can show that there exists a Markov Perfect Equilibrium such that in each stage the voters' rule strategies are given by (1) and the politicians' efforts are given by (2).

1.2.5 Dynamics

In this section I calculate the equilibrium probabilities of transition between state S (where the politicians are members of the same party) and state D (where the politicians belong

to different parties). I denote by P_{kl} the probability that a city in state k will shift to state l in the next period, $k, l \in \{S, D\}$. I establish the following result.

Proposition 3 *The matrix of the equilibrium one-step transition probabilities between states S and D is*

$$\begin{bmatrix} P_{SS} & P_{SD} \\ P_{DS} & P_{DD} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} + \frac{1}{\pi} \arctan \frac{\lambda_M^S + \lambda_G^S}{1 - \lambda_M^S \lambda_G^S} & \frac{1}{2} - \frac{1}{\pi} \arctan \frac{\lambda_M^S + \lambda_G^S}{1 - \lambda_M^S \lambda_G^S} \\ \frac{1}{2} + \frac{1}{\pi} \arctan \frac{\lambda_M^D + \lambda_G^D}{1 - \lambda_M^D \lambda_G^D} & \frac{1}{2} - \frac{1}{\pi} \arctan \frac{\lambda_M^D + \lambda_G^D}{1 - \lambda_M^D \lambda_G^D} \end{bmatrix}$$

where $\arctan(\cdot)$ is the arctangent function.

Note that independently of the current state, the next state is more likely to be S than D . Indeed, the probability of the next state being a split-ticket state is never greater than $\frac{1}{2}$: $P_{.D} \in [0, \frac{1}{2}]$. The intuition for this result is as follows. If the politicians currently belong to the same party (state S), the voters adopt a joint rule under which the reelection outcomes are positively correlated: the incumbents are more likely to be reelected together or dismissed together than they are to receive opposite rewards. Thus, the next state is more likely to be S . If the politicians are currently members of different parties (state D), then the voters use a joint rule under which the reelection outcomes are negatively correlated. Thus, it is more likely that one incumbent will be dismissed while the other is reelected, and again the next state is more likely to be S than D . To confirm this intuition, in Figure 3 I depict the politicians' reelection outcomes under equilibrium rules β_M^* and β_G^* in the two-dimensional space of performances p_M and p_G . The density function of the joint distribution of p_M and p_G is symmetric around (a_M^*, a_G^*) .

Furthermore, independently of the current state, the probability $P_{.D}$ that the next state will be D is decreasing in λ_M and λ_G . This probability takes its minimal value of 0 when $\lambda_M = \lambda_G = 1$, and its maximal value of $\frac{1}{2}$ when $\lambda_M = \lambda_G = 0$. Intuitively, the more loyal politicians are to their parties, the more correlated (positively if S or negatively if D) the optimal performance evaluation rules. The outcome S is more probable for both current states, as explained above, so stronger party loyalty just increases the probability of this outcome.

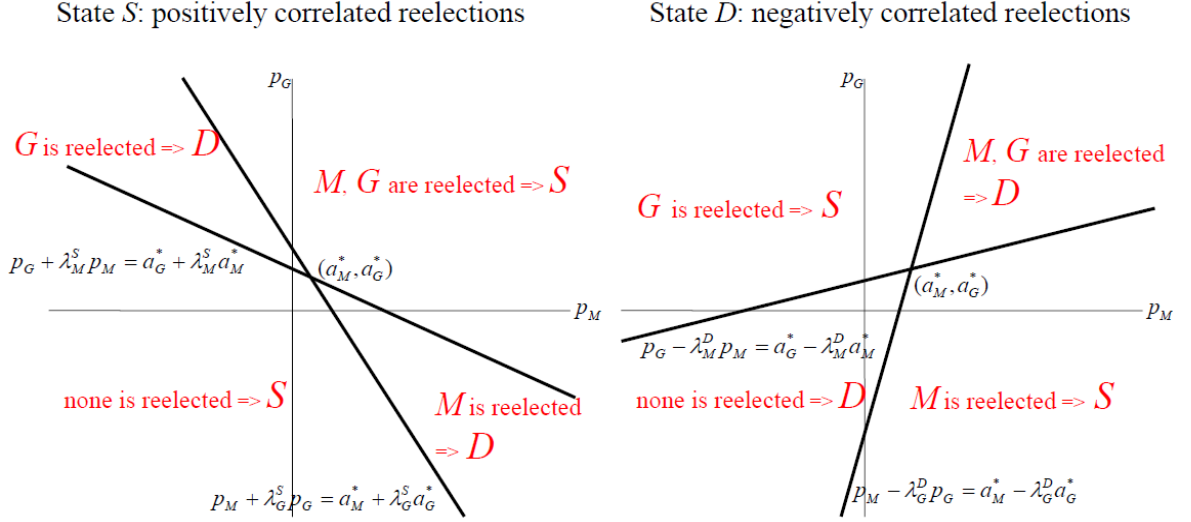


Figure 3: Politicians' reelection outcomes under equilibrium rules β_M^* and β_G^* in states S and D .

How does the current state affect the probability that the next state is state k , $k \in \{S, D\}$? First, if politicians prefer incumbents ($\lambda_i^S \geq \lambda_i^D$) then $P_{DD} \geq P_{SD}$ and $P_{SS} \geq P_{DS}$; although S is always more likely than D , the next state is more likely to be k if the current state is k . Intuitively, due to the politicians' preference for incumbents, the politicians' incentives are more correlated in state S than in state D . The voters are therefore more likely to adopt positively correlated reelection rules in S than negatively correlated rules in D . While both states are more likely to shift to S in the next period, this S outcome is more likely to occur if the current state is S . By the same logic, a D outcome is more likely if the current state is D than if the current state is S . Second, if politicians prefer newcomers ($\lambda_i^S < \lambda_i^D$) then $P_{DD} < P_{SD}$ and $P_{SS} < P_{DS}$. Here the politicians' incentives are more correlated in state D than in state S . Thus, the joint reelection rules are more correlated (in absolute value) in state D than in state S . So state S is more likely to occur if the current state is D , while state D is more likely to occur if the current state is S .

Recall that state S or state D occurs when the voters did not split or split tickets respectively. I conclude that in this simultaneous elections framework, ticket splitting is less likely than voting for candidates from the same party (i.e., $P_D \leq P_S$). Moreover, the probability of split-ticket voting depends on ticket splitting in the previous period. If

politicians prefer incumbents then the voters are more likely to split tickets if in the previous period they also split tickets, i.e.,

$$P_{DD} \geq P_{SD} \text{ if } \lambda_M^S \geq \lambda_M^D \text{ and } \lambda_G^S \geq \lambda_G^D.$$

If politicians prefer newcomers then the voters are more likely to split tickets if in the previous period they did not split ticket, i.e.,

$$P_{DD} < P_{SD} \text{ if } \lambda_M^S < \lambda_M^D \text{ and } \lambda_G^S < \lambda_G^D.$$

1.3 *Ticket Splitting in a Region with Small Municipalities*

In this section I assume that a region consists of n municipalities, and that each is pivotal for the outcome of the regional election with some small probability. The main insights and intuitions of the large city case (Section 1.2) do not change qualitatively. Still, some novel results arise. This section stresses the novel assumptions and results, referring to the previous analysis whenever appropriate.

The model is identical to that presented in Section 1.2, except for the following changes. The region consists of n municipalities with population shares ρ_i , $\sum_{i=1}^n \rho_i = 1$. At the beginning of each period the voters in municipality i elect mayor M_i and vote for governor G in the simultaneous elections. The probability that municipality i is pivotal for the outcome of the regional election is equal to its population share ρ_i . While in office, mayors M_i and governor G implement policies which are determined by their unobservable efforts a_i and a_G respectively. The performances p_i and p_G are observed with independent and unobservable noises ε_i and ε_G respectively:

$$p_i = a_i + \varepsilon_i$$

$$p_G = a_G + \varepsilon_G$$

with $\varepsilon_i, \varepsilon_G \sim N(0, \sigma^2)$, $i = 1, \dots, n$.

Mayor M_i chooses effort a_i to maximize her utility, given by

$$\Pi_i(a_i, a_G) - \frac{a_i^2}{2} = \begin{cases} \Pr_i(a_i, a_G) + \lambda_i^S \Pr_G(a_i, a_G) - \frac{a_i^2}{2} & \text{if } S_i \\ \Pr_i(a_i, a_G) + \lambda_i^D (1 - \Pr_G(a_i, a_G)) - \frac{a_i^2}{2} & \text{if } D_i, \end{cases}$$

where S_i and D_i denote the states where mayor M_i and governor G are members of the same party or different parties as before. This utility function implies that the mayor is office-motivated and prefers the governor to be a politician from the same party. The results do not change if I extend the mayor's partisan preferences to include all other mayors in the region.

Governor G is also office-motivated and loyal to her political party. She prefers to see members of her own party in all the offices M_1, \dots, M_n . However, I assume that a governor cares more about her party's election chances in larger cities. In other words, the larger the population of the city, the more the governor wants its mayor to be a member of her own political party. This assumption reflects the idea that party affiliation might be of less importance in smaller municipalities, either because candidates are more likely to run as independents (or be affiliated with a minor party) or because the voters have personal knowledge of the candidates for mayor office. Recall that one of the major roles of political parties is to provide information about unknown politicians. Formally,

$$\begin{aligned} \Pi_G(a_1, \dots, a_n, a_G) - \frac{a_G^2}{2} &= Pr_G(a_1, \dots, a_n, a_G) + \lambda_G^S \sum_{i=1}^n \rho_i I_i Pr_i(a_i, a_G) + \\ &\quad \lambda_G^D \sum_{i=1}^n \rho_i (1 - I_i) (1 - Pr_i(a_i, a_G)) - \frac{a_G^2}{2}. \end{aligned}$$

I_i is the indicator function of state S_i , defined as

$$I_i = \begin{cases} 1 & \text{if } S_i \\ 0 & \text{if } D_i. \end{cases}$$

The voters in municipality i care about the politicians' performances in each period according to a linear utility function

$$p_i + p_G,$$

where p_i and p_G are observable at the end of each period. The voters coordinate in choosing a linear performance evaluation rule (β_i, b_i) to reward mayor M_i , and a linear rule (β_i^G, b_i^G) to reward governor G . The probability that mayor M_i is reelected equals

$$Pr_i(a_i, a_G) = P(\{p_i(a_i) + \beta_i p_G(a_G) \geq b_i\}).$$

As for the governor, I assume that each municipality i has a probability equal to its population share ρ_i to be pivotal in the regional election. The probability that governor G is reelected is therefore additively separable, and equal to a weighted sum of the probabilities of getting a majority in each municipality. Each municipality's term is weighted with its population share:

$$Pr_G(a_1, \dots, a_n, a_G) = \sum_{i=1}^n \rho_i P(\{p_G(a_G) + \beta_i^G p_i(a_i) \geq b_i^G\}).$$

I skip the equilibrium definitions and the discussion, which are analogous to the large city model in Section 1.2. Next, I characterize the equilibrium in rule strategies in the case of n small municipalities.

Proposition 4 *There exists an equilibrium in rule strategies $(\beta_i^*, \beta_i^{G*})$ $i = 1, \dots, n$, given by*

$$\beta_i^* = \begin{cases} \rho_i \lambda_G^S & \text{if } S_i \\ -\rho_i \lambda_G^D & \text{if } D_i \end{cases} \quad \text{and} \quad \beta_i^{G*} = \begin{cases} \lambda_i^S & \text{if } S_i \\ -\lambda_i^D & \text{if } D_i. \end{cases}$$

The politicians' equilibrium efforts a_i^, a_G^* are equal to*

$$a_i^* = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \left(\frac{1}{\sqrt{1+(\rho_i \lambda_G^S)^2}} + \frac{\rho_i (\lambda_i^S)^2}{\sqrt{1+(\lambda_i^S)^2}} \right) & \text{if } S_i \\ \frac{1}{\sqrt{2\pi}\sigma} \left(\frac{1}{\sqrt{1+(\rho_i \lambda_G^D)^2}} + \frac{\rho_i (\lambda_i^D)^2}{\sqrt{1+(\lambda_i^D)^2}} \right) & \text{if } D_i \end{cases}$$

$$a_G^* = \frac{1}{\sqrt{2\pi}\sigma} \sum_{i=1}^n \rho_i \left(\frac{1}{\sqrt{1+(\lambda_i^S)^2 I_i + (\lambda_i^D)^2 (1-I_i)}} + \frac{\rho_i ((\lambda_G^S)^2 I_i + (\lambda_G^D)^2 (1-I_i))}{\sqrt{1+(\rho_i \lambda_G^S)^2 I_i + (\rho_i \lambda_G^D)^2 (1-I_i)}} \right).$$

The equilibrium analysis and intuition for the case of n small municipalities do not differ qualitatively from the large city case presented in Section 1.2. The only new prediction is that the correlation (positive in state S_i or negative in state D_i) between the optimal reward rules for mayor M_i and governor G is stronger in large cities than in small cities. Intuitively, the larger the municipality, the more the governor wants its mayor to belong to the same party, so the more correlated the politicians' incentives. As a result, the more the voters correlate (positively if S_i or negatively if D_i) the optimal reward rule for mayor M_i with performance of governor G to motivate the latter to exert higher effort.

I turn now to a dynamic analysis of ticket splitting in small municipalities, which reveals more novel insights. Consider a municipality that splits tickets in the current period, i.e., the voters elect a mayor from one party and vote for a governor from the other party. This does not necessarily imply the election of the city's preferred governor, since each municipality has only a small probability of swaying the regional election. Thus, ticket splitting in municipality i does not necessarily result in the state D_i where mayor and governor are members of different parties. Note that in the large city model, ticket splitting always leads to state D .

I find the equilibrium probabilities of transition between the ticket-splitting and non ticket-splitting states in municipality i . I denote by Y_i the state where voters in municipality i split tickets (Y stands for "yes"), and by N_i the state where voters in municipality i do not split tickets (N stands for "no"). I denote by q_i the probability that the governor who wins a majority in municipality i is actually elected. In other words, probability q_i is equal to the probability that an incumbent governor gets the majority in municipality i and is also reelected, plus the probability that an incumbent governor does not get the majority in municipality i and is not reelected. I now establish the following result.

Proposition 5 *The matrix of the equilibrium one-step transition probabilities between states N_i and Y_i is*

$$\begin{bmatrix} P_{N_i N_i} & P_{N_i Y_i} \\ P_{Y_i N_i} & P_{Y_i Y_i} \end{bmatrix}$$

where

$$\begin{aligned} P_{N_i Y_i} &= \frac{1}{2} - \frac{1}{\pi} \left(q_i \arctan \frac{\lambda_i^S + \rho_i \lambda_G^S}{1 - \rho_i \lambda_i^S \lambda_G^S} + (1 - q_i) \arctan \frac{\lambda_i^D + \rho_i \lambda_G^D}{1 - \rho_i \lambda_i^D \lambda_G^D} \right) \\ P_{Y_i Y_i} &= \frac{1}{2} - \frac{1}{\pi} \left(q_i \arctan \frac{\lambda_i^D + \rho_i \lambda_G^D}{1 - \rho_i \lambda_i^D \lambda_G^D} + (1 - q_i) \arctan \frac{\lambda_i^S + \rho_i \lambda_G^S}{1 - \rho_i \lambda_i^S \lambda_G^S} \right) \end{aligned}$$

with

$$q_i = 1 - \frac{1}{\pi} \sum_{j \neq i} \rho_j \arctan \sqrt{(\beta_i^{G*})^2 + (\beta_j^{G*})^2 + (\beta_i^{G*} \beta_j^{G*})^2}.$$

Proposition 5 generalizes the dynamic predictions of the large city case studied in Section 1.2. In analogy with previous results, the next state is more likely to be non ticket-splitting

state N_i than ticket-splitting state Y_i regardless of the current state, since $P_{Y_i} \leq \frac{1}{2}$. Furthermore, if politicians prefer incumbents ($\lambda_i^S \geq \lambda_i^D$ and $\lambda_G^S \geq \lambda_G^D$) then the municipality is more likely to split tickets if in the previous period it also split tickets ($P_{Y_i Y_i} \geq P_{N_i Y_i}$). In case politicians prefer newcomers ($\lambda_i^S < \lambda_i^D$ and $\lambda_G^S < \lambda_G^D$) then the municipality is more likely to split tickets if in the previous period it did not split tickets ($P_{Y_i Y_i} < P_{N_i Y_i}$).

The following result arises only in the small-city model. According to Proposition 5, ticket splitting is more likely to happen in small municipalities than in large ones regardless of the current state. The probabilities $P_{N_i Y_i}$ and $P_{Y_i Y_i}$ are decreasing functions of the population share ρ_i . Intuitively, a governor cares less about the party affiliation of small-town mayors. As the politicians' incentives are less correlated in small municipalities than in large ones, the voters adopt less correlated joint performance evaluation rules. As a result, the incumbents are more likely to be evaluated according to their individual performance, which increases the probability of ticket splitting.

I summarize these findings in the following corollary.

Corollary 6 *1. Regardless of the current state, ticket splitting is less likely than voting for candidates from the same party.*

2. The probability of split-ticket voting depends on ticket splitting in the previous period.

3. Ticket splitting is more likely in small municipalities than in large ones.

In the next section I show that the predictions of Corollary 6 are consistent with ticket splitting patterns in Spain.

1.4 Empirical Model

The goal of this section is to estimate the probability of ticket splitting and show that the predictions of Corollary 6 are consistent with the patterns of ticket splitting observed in Spain. I use an unbalanced panel data set on simultaneous municipal and regional elections held in Spain in 1983, 1987, 1991, 1995, 1999, 2003 and 2007.¹⁰

¹⁰One might be concerned whether the majoritarian model presented here is applicable to the empirical context, since Spanish regional elections use a proportional representation system. In response, I stress that

I apply the binary response model, which employs a probit link function. A logit model yields the same qualitative results. However, the probit model is more consistent with my theoretical framework, where reelections are determined by normally distributed noise.¹¹

Denote by y_{it} a binary variable that takes value 1 if municipality i splits tickets in period t , and 0 otherwise. Furthermore, ρ_i stands for municipality i 's population share, x_{it} for control variables, ξ_r for region effects (where r denotes the region that municipality i belongs to) and ξ_t for year effects. The theoretical model suggests the following structure for the estimating equation:

$$P(y_{it} = 1 | \rho_i, y_{it-1}, x_{it}, \xi_r, \xi_t) = \Phi(\mu_0 + \mu_1 \rho_i + \mu_2 y_{it-1} + \mu_3 x_{it} + \xi_r + \xi_t),$$

where $\Phi(\cdot)$ denotes the cumulative function of the standard normal distribution. The region effects and year effects are included as dummies.

Since the specification does not include time invariant municipality effects, I can estimate the model by pooling all the cross sections. Otherwise, in order to obtain consistent estimates I would have to apply either random effects or fixed effects estimation methods. This would not be straightforward, given the non-linear nature of the model.

Before proceeding, I will discuss what effects are expected in the empirical analysis. According to Corollary 6 ticket splitting is more likely in smaller municipalities, which suggests $\mu_1 < 0$. The effect of ticket splitting in the previous period should be reflected by $\mu_2 \neq 0$. Finally, I hypothesize that the predicted probability of ticket splitting is less than 0.5: $\hat{P}(y_{it} = 1) < 0.5$.

1.4.1 Data Description

In Spain, local municipal and regional elections occur simultaneously every four years in 13 out of 17 regions (the so-called autonomous communities).¹² The two leading parties are

in Spain, the regional leader of the party that gets the most seats in the regional Legislative Assembly is usually elected as a president of the corresponding autonomous community. As such, ordinary citizens often regard the regional election as a way of determining the president of their autonomous community, rather than an election of their representatives.

¹¹The maximum likelihood estimation yields the same qualitative results.

¹²Municipal and regional elections take place simultaneously in Aragon, Principality of Asturias, Balearic Islands, Canary Islands, Cantabria, Castile-La Mancha, Castile and León, Extremadura, La Rioja, Community of Madrid, Region of Murcia, Foral Community of Navarre and Valencian Community. In Andalusia, Basque Country, Catalonia and Galicia, municipal and regional elections are held on different dates.

Partido Popular (PP) and Partido Socialista Obrero Español (PSOE).¹³ To build the data set I use the aggregate election results of 10 Spanish regions, which are partially available online at the official websites of the regional governments and the Spanish Ministry of the Interior.¹⁴

The sample consists of 3218 municipalities, and depending on the region, covers from 4 to 7 election years from 1983 to 2007. Initially, each observation (of municipality i in election year t) includes a census, the number of abstainers, the votes for PP, the votes for PSOE, and the votes for other parties in both municipal and regional elections.

In the theoretical analysis I assumed that all voters participate in both municipal and regional elections. To meet this requirement, from the initial sample I discard all observations where the number of voters in municipal elections differs significantly from the number of voters in regional elections (the maximum allowable difference is 5%). This ensures that almost the same electorate participated in both elections. Next, I exclude all observations where a third party obtained more votes than either PP or PSOE, in either the municipal or regional elections. All observations thus have the same two leading parties, in line with the theoretical model. Then I define the binary variable y_{it} such that $y_{it} = 1$ if different parties obtained the largest number of votes in the municipal and regional elections (ticket splitting) and $y_{it} = 0$ if the same party received the largest number of votes in both elections (no ticket splitting).

I use the census share of municipality i in a region during the last observed election year as a proxy for the population share ρ_i . The per capita GDP by province (in thousands of euros) serves as the control variable x_{it} .¹⁵ Table 3 and Figure 4 in the Appendix provide descriptive statistics and characteristics of the final sample.

¹³There are also several minor parties; for example, Izquierda Unida has considerable support in some regions.

¹⁴Some data were kindly provided by the statistical institutes of the corresponding regions, and are available upon request. The community of Castile and León is not included in my analysis because the data on regional elections in this community are not available. The Canary Islands and the Foral Community of Navarre are not included in the data set, because local parties apart from PP and PSOE enjoy widespread support in these regions and the theoretical model assumes just two political parties.

¹⁵In Spain, provinces are administrative subdivisions of autonomous communities. In turn, municipalities are subdivisions of provinces.

	$P(\text{ticket splitting in } i \text{ at } t) = P(y_{it} = 1)$		
	(1)	(2)	(3)
Pop. share, μ_1	-2.544* (1.332)	-2.616** (1.309)	-2.793* (1.456)
TS in $t - 1$, μ_2	0.734*** (0.043)	0.734*** (0.043)	0.741*** (0.04)
Observations	4183	4183	4177

Robust standard errors clustered by region in parentheses. Significant at 10%-*; 5%-**; 1%-***
(1) – only region dummies; (2) – region dummies and year dummies; (3) – region-year dummies.

Table 1: Ticket splitting (TS) and municipality population share.

	obs.	mean	std. dev.	min	max	std. err.	95% conf. interval
$\hat{P}(y_{it} = 1)$	4183	0.193	0.106	0.006	0.487	0.002	0.190 0.196

Table 2: Summary statistics of the predicted probability of ticket splitting $\hat{P}(y_{it} = 1)$.

1.4.2 Empirical Results

Table 1 presents the coefficients of interest in the panel regressions. The first regression includes only region dummies. The second regression includes both region and year dummies. In the third regression, the region dummies are interacted with each year dummy.

First, consider coefficient μ_1 for the population share. It is significantly negative in all specifications. I conclude that ticket splitting is more likely to occur in smaller municipalities, as predicted. Second, coefficient μ_2 for the effect of ticket splitting in the previous period is significantly positive in all specifications. This result shows that in the sample ticket splitting is more likely in municipalities where the voters split tickets in the previous period.

Finally, Table 2 presents summary statistics of the predicted probability of ticket splitting $\hat{P}(y_{it} = 1)$ in the second regression, which includes both region and year dummies. The maximal predicted probability does not exceed 0.5 and is equal to 0.487. So I conclude that ticket splitting is less likely than partisan voting.

In sum, the predictions of the theoretical model, which are summarized in Corollary 6, are consistent with these empirical findings on the dynamics of simultaneous two-party elections.

1.5 Conclusion

This paper applies an implicit incentive approach to study split-ticket voting in simultaneous municipal and regional elections. In a political agency model with moral hazard, ticket splitting is a natural outcome of the optimal implicit reward schemes that voters use to motivate the politicians' efforts.

I assume that incentives of local and regional politicians are correlated, as a mayor/governor prefers her counterpart (the governor/mayor) to be affiliated with the same political party. Voters thus are better off adopting a joint performance evaluation rule rather than individual performance evaluation rules to reward the incumbents. These rules affect the dynamics of split-ticket voting. In particular, the model suggests that ticket splitting is less likely than voting for candidates from the same party. Moreover, the probability of split-ticket voting depends on ticket-splitting in the previous period. Finally, ticket splitting is more likely in smaller municipalities where the party affiliation of a mayor is assumed to be of less importance to the governor. These results are consistent with empirical evidence obtained from Spanish elections where voters are more likely to split tickets if in the previous period they also did so.

I have focused on single task policies. However, in reality public policies pursue many goals. So it is of interest to study split-ticket voting under a more realistic assumption of multiple-task policy where the problem of effort allocation among tasks can create policy trade-offs. To refine the empirical results, one could apply other estimation methods such as a Markov switching model. It would also be interesting to examine data from other countries where municipal and regional elections are held simultaneously. I leave these tasks for future research.

1.6 Appendix

Throughout the Appendix, I use F to denote the normal distribution function and f for the corresponding density.

Proof of Proposition 1. Under linear performance evaluation rules (β_i, b_i) the probability of being reelected for office i is

$$Pr_i(a_i, a_j) = P(\{\varepsilon_i + \beta_i \varepsilon_j \geq b_i - a_i - \beta_i a_j\}) = 1 - F_{\varepsilon_i + \beta_i \varepsilon_j}(b_i - a_i - \beta_i a_j),$$

where noises ε_i and ε_j ($i, j \in \{M, G\}$, $i \neq j$) are independent normally distributed random variables, so by the convolution formula $\varepsilon_i + \beta_i \varepsilon_j \sim N(0, (1 + \beta_i^2) \sigma^2)$. Politician i 's utility is

$$\Pi_i(a_i, a_j) - \frac{a_i^2}{2} = \begin{cases} 1 - F_{\varepsilon_i + \beta_i \varepsilon_j}(b_i - a_i - \beta_i a_j) + \lambda_i^S \left(1 - F_{\varepsilon_j + \beta_j \varepsilon_i}(b_j - a_j - \beta_j a_i)\right) - \frac{a_i^2}{2} & \text{if } S \\ 1 - F_{\varepsilon_i + \beta_i \varepsilon_j}(b_i - a_i - \beta_i a_j) + \lambda_i^D F_{\varepsilon_j + \beta_j \varepsilon_i}(b_j - a_j - \beta_j a_i) - \frac{a_i^2}{2} & \text{if } D. \end{cases}$$

The first-order conditions with respect to actual effort a_i , taking $b_i = a_i^e + \beta_i a_j^e$ and $b_j = a_j^e + \beta_j a_i^e$ as given, are

$$\begin{cases} f_{\varepsilon_i + \beta_i \varepsilon_j}(b_i - a_i - \beta_i a_j) + \lambda_i^S \beta_j f_{\varepsilon_j + \beta_j \varepsilon_i}(b_j - a_j - \beta_j a_i) - a_i = 0 & \text{if } S \\ f_{\varepsilon_i + \beta_i \varepsilon_j}(b_i - a_i - \beta_i a_j) - \lambda_i^D \beta_j f_{\varepsilon_j + \beta_j \varepsilon_i}(b_j - a_j - \beta_j a_i) - a_i = 0 & \text{if } D. \end{cases}$$

Imposing the equilibrium requirement $a_i = a_i^e$ yields

$$a_i^e = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} \left(\frac{1}{\sqrt{1+\beta_i^2}} + \frac{\lambda_i^S \beta_j}{\sqrt{1+\beta_j^2}} \right) & \text{if } S \\ \frac{1}{\sqrt{2\pi}\sigma} \left(\frac{1}{\sqrt{1+\beta_i^2}} - \frac{\lambda_i^D \beta_j}{\sqrt{1+\beta_j^2}} \right) & \text{if } D, \end{cases}$$

which completes the proof. ■

Proof of Proposition 3. The reelection of incumbent i is determined by random variable $\varepsilon_i + \beta_i \varepsilon_j \sim N(0, (1 + \beta_i^2) \sigma^2)$, $i, j \in \{M, G\}$, $i \neq j$. The density function of bivariate normal distribution of random variables $\varepsilon_M + \beta_M \varepsilon_G$ and $\varepsilon_G + \beta_G \varepsilon_M$, denoted by $f_{\varepsilon_M + \beta_M \varepsilon_G, \varepsilon_G + \beta_G \varepsilon_M}(x, y)$, is

$$f_{\varepsilon_M + \beta_M \varepsilon_G, \varepsilon_G + \beta_G \varepsilon_M}(x, y) = \frac{1}{2\pi\sigma^2 \sqrt{(\beta_M \beta_G - 1)^2}} \exp \left\{ -\frac{(x - y\beta_M)^2 + (y - x\beta_G)^2}{2\sigma^2 (\beta_M \beta_G - 1)^2} \right\}.$$

The transition from state k back to state k , $k \in \{S, D\}$, occurs either when both incumbents are reappointed or when none of them is reappointed (so, opponents from rival parties are

elected). Denote by $p_i^* = a_i^* + \varepsilon_i$ the performance of politician i in equilibrium. The equilibrium transition probabilities are

$$\begin{aligned}
P_{SS} &= P(\{p_M^* + \lambda_G^S p_G^* \geq a_M^* + \lambda_G^S a_G^*\} \cap \{p_G^* + \lambda_M^S p_M^* \geq a_G^* + \lambda_M^S a_M^*\}) + \\
&P(\{p_M^* + \lambda_G^S p_G^* < a_M^* + \lambda_G^S a_G^*\} \cap \{p_G^* + \lambda_M^S p_M^* < a_G^* + \lambda_M^S a_M^*\}) = \\
&P(\{\varepsilon_M + \lambda_G^S \varepsilon_G \geq 0\} \cap \{\varepsilon_G + \lambda_M^S \varepsilon_M \geq 0\}) + \\
&P(\{\varepsilon_M + \lambda_G^S \varepsilon_G < 0\} \cap \{\varepsilon_G + \lambda_M^S \varepsilon_M < 0\}) = \\
&\int_0^{+\infty} \int_0^{+\infty} f_{\varepsilon_M + \lambda_G^S \varepsilon_G, \varepsilon_G + \lambda_M^S \varepsilon_M}(x, y) dx dy + \int_{-\infty}^0 \int_{-\infty}^0 f_{\varepsilon_M + \lambda_G^S \varepsilon_G, \varepsilon_G + \lambda_M^S \varepsilon_M}(x, y) dx dy = \\
&\frac{1}{2} + \frac{1}{\pi} \arctan \frac{\lambda_M^S + \lambda_G^S}{1 - \lambda_M^S \lambda_G^S}
\end{aligned}$$

$$\begin{aligned}
P_{DD} &= P(\{p_M^* - \lambda_G^D p_G^* \geq a_M^* - \lambda_G^D a_G^*\} \cap \{p_G^* - \lambda_M^D p_M^* \geq a_G^* - \lambda_M^D a_M^*\}) + \\
&P(\{p_M^* - \lambda_G^D p_G^* < a_M^* - \lambda_G^D a_G^*\} \cap \{p_G^* - \lambda_M^D p_M^* < a_G^* - \lambda_M^D a_M^*\}) = \\
&P(\{\varepsilon_M - \lambda_G^D \varepsilon_G \geq 0\} \cap \{\varepsilon_G - \lambda_M^D \varepsilon_M \geq 0\}) + \\
&P(\{\varepsilon_M - \lambda_G^D \varepsilon_G < 0\} \cap \{\varepsilon_G - \lambda_M^D \varepsilon_M < 0\}) = \\
&\int_0^{+\infty} \int_0^{+\infty} f_{\varepsilon_M - \lambda_G^D \varepsilon_G, \varepsilon_G - \lambda_M^D \varepsilon_M}(x, y) dx dy + \int_{-\infty}^0 \int_{-\infty}^0 f_{\varepsilon_M - \lambda_G^D \varepsilon_G, \varepsilon_G - \lambda_M^D \varepsilon_M}(x, y) dx dy = \\
&\frac{1}{2} - \frac{1}{\pi} \arctan \frac{\lambda_M^D + \lambda_G^D}{1 - \lambda_M^D \lambda_G^D}
\end{aligned}$$

$$P_{SD} = 1 - P_{SS} = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{\lambda_M^S + \lambda_G^S}{1 - \lambda_M^S \lambda_G^S}$$

$$P_{DS} = 1 - P_{DD} = \frac{1}{2} + \frac{1}{\pi} \arctan \frac{\lambda_M^D + \lambda_G^D}{1 - \lambda_M^D \lambda_G^D},$$

where $\arctan(\cdot)$ is an arctangent function. ■

Proof of Proposition 4. Under linear performance evaluation rules (β_i, b_i) and (β_i^G, b_i^G) , $i = 1, \dots, n$, mayor M_i 's utility is

$$\begin{aligned}
&\Pi_i(a_i, a_G) - \frac{a_i^2}{2} = \\
&\begin{cases} 1 - F_{\varepsilon_i + \beta_i \varepsilon_G}(b_i - a_i - \beta_i a_G) + \lambda_i^S \sum_{j=1}^n \rho_j \left(1 - F_{\varepsilon_G + \beta_j^G \varepsilon_j}(b_j^G - a_G - \beta_j^G a_j)\right) - \frac{a_i^2}{2} & \text{if } S_i \\ 1 - F_{\varepsilon_i + \beta_i \varepsilon_G}(b_i - a_i - \beta_i a_G) + \lambda_i^D \sum_{j=1}^n \rho_j F_{\varepsilon_G + \beta_j^G \varepsilon_j}(b_j^G - a_G - \beta_j^G a_j) - \frac{a_i^2}{2} & \text{if } D_i. \end{cases}
\end{aligned}$$

The first-order conditions with respect to actual effort a_i , taking $b_i = a_i^e + \beta_i a_G^e$ and $b_j^G = a_G^e + \beta_j^G a_j^e$ as given, are

$$\begin{cases} f_{\varepsilon_i + \beta_i \varepsilon_G}(b_i - a_i - \beta_i a_G) + \lambda_i^S \rho_i \beta_i^G f_{\varepsilon_G + \beta_i^G \varepsilon_i}(b_j^G - a_G - \beta_j^G a_j) - a_i = 0 & \text{if } S_i \\ f_{\varepsilon_i + \beta_i \varepsilon_G}(b_i - a_i - \beta_i a_G) - \lambda_i^D \rho_i \beta_i^G f_{\varepsilon_G + \beta_i^G \varepsilon_i}(b_j^G - a_G - \beta_j^G a_j) - a_i = 0 & \text{if } D_i. \end{cases}$$

I impose the equilibrium requirements $a_i = a_i^e$ and $a_G = a_G^e$ to get mayor M_i 's equilibrium effort strategy

$$a_i^e = \begin{cases} \frac{1}{\sqrt{2\pi\sigma}} \left(\frac{1}{\sqrt{1+\beta_i^2}} + \frac{\lambda_i^S \rho_i \beta_i^G}{\sqrt{1+\beta_i^{G2}}} \right) & \text{if } S_i \\ \frac{1}{\sqrt{2\pi\sigma}} \left(\frac{1}{\sqrt{1+\beta_i^2}} - \frac{\lambda_i^D \rho_i \beta_i^G}{\sqrt{1+\beta_i^{G2}}} \right) & \text{if } D_i. \end{cases}$$

Next, consider governor G 's utility:

$$\begin{aligned} \Pi_G(a_1, \dots, a_n, a_G) - \frac{a_G^2}{2} &= \sum_{j=1}^n \rho_j \left(1 - F_{\varepsilon_G + \beta_j^G \varepsilon_j}(b_j^G - a_G - \beta_j^G a_j) \right) + \\ \lambda_G^S \sum_{j=1}^n I_j \rho_j \left(1 - F_{\varepsilon_j + \beta_j \varepsilon_G}(b_j - a_j - \beta_j a_G) \right) &+ \lambda_G^D \sum_{j=1}^n (1 - I_j) \rho_j F_{\varepsilon_j + \beta_j \varepsilon_G}(b_j - a_j - \beta_j a_G) - \frac{a_G^2}{2}. \end{aligned}$$

The first-order condition with respect to actual effort a_G , taking $b_j = a_j^e + \beta_j a_G^e$ and $b_j^G = a_G^e + \beta_j^G a_j^e$ as given, yields

$$\begin{aligned} \sum_{j=1}^n \rho_j f_{\varepsilon_G + \beta_j^G \varepsilon_j}(b_j^G - a_G - \beta_j^G a_j) + \lambda_G^S \sum_{j=1}^n I_j \rho_j \beta_j f_{\varepsilon_j + \beta_j \varepsilon_G}(b_j - a_j - \beta_j a_G) - \\ \lambda_G^D \sum_{j=1}^n (1 - I_j) \rho_j \beta_j f_{\varepsilon_j + \beta_j \varepsilon_G}(b_j - a_j - \beta_j a_G) - a_G = 0. \end{aligned}$$

Imposing the equilibrium requirements $a_j = a_j^e$ ($j = 1, \dots, n$) and $a_G = a_G^e$ yields governor G 's equilibrium effort strategy

$$a_G^e = \frac{1}{\sqrt{2\pi\sigma}} \sum_{j=1}^n \rho_j \left(\frac{1}{\sqrt{1+\beta_j^{G2}}} + \frac{\beta_j (\lambda_G^S I_j - \lambda_G^D (1 - I_j))}{\sqrt{1+\beta_j^2}} \right).$$

Finally, maximizing $a_i^e + a_G^e$ with respect to β_i and β_i^G yields an equilibrium in rule strategies $(\beta_i^*, \beta_i^{G*})$ and politicians' equilibrium efforts $a_i^* = a_i^e(\beta_i^*, \beta_i^{G*})$ and $a_G^* = a_G^e(\beta_i^*, \beta_i^{G*})$, which completes the proof. ■

Proof of Proposition 5. The matrix of the equilibrium one-step transition probabilities between states N_i and Y_i equals

$$\begin{bmatrix} P_{N_i N_i} & P_{N_i Y_i} \\ P_{Y_i N_i} & P_{Y_i Y_i} \end{bmatrix} = \begin{bmatrix} q_i P_{S_i N_i} + (1 - q_i) P_{D_i N_i} & q_i P_{S_i Y_i} + (1 - q_i) P_{D_i Y_i} \\ q_i P_{D_i N_i} + (1 - q_i) P_{S_i N_i} & q_i P_{D_i Y_i} + (1 - q_i) P_{S_i Y_i} \end{bmatrix}$$

Refer to the proof of Proposition 3 to find transition probabilities to states N_i and Y_i from states S_i and D_i :

$$P_{S_i N_i} = \frac{1}{2} + \frac{1}{\pi} \arctan \frac{\lambda_i^S + \rho_i \lambda_G^S}{1 - \rho_i \lambda_i^S \lambda_G^S} \quad \text{and} \quad P_{S_i Y_i} = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{\lambda_i^S + \rho_i \lambda_G^S}{1 - \rho_i \lambda_i^S \lambda_G^S}$$

$$P_{D_i N_i} = \frac{1}{2} + \frac{1}{\pi} \arctan \frac{\lambda_i^D + \rho_i \lambda_G^D}{1 - \rho_i \lambda_i^D \lambda_G^D} \quad \text{and} \quad P_{D_i Y_i} = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{\lambda_i^D + \rho_i \lambda_G^D}{1 - \rho_i \lambda_i^D \lambda_G^D}.$$

Next,

$$\begin{aligned} q_i &= P(\{G \text{ gets majority in } i\} \cap \{G \text{ is reelected}\}) + \\ &P(\{G \text{ does not get majority in } i\} \cap \{G \text{ is not reelected}\}) = \\ &2 \sum_{j=1}^n \rho_j P(\{p_G^* + \beta_i^{G*} p_i^* \geq a_G^* + \beta_i^{G*} a_i^*\} \cap \{p_G^* + \beta_j^{G*} p_j^* \geq a_G^* + \beta_j^{G*} a_j^*\}) = \\ &2 \sum_{j=1}^n \rho_j P(\{\varepsilon_G + \beta_i^{G*} \varepsilon_i \geq 0\} \cap \{\varepsilon_G + \beta_j^{G*} \varepsilon_j \geq 0\}) = \\ &\rho_i + 2 \sum_{j \neq i} \rho_j \int_0^{+\infty} \int_0^{+\infty} f_{\varepsilon_G + \beta_i^{G*} \varepsilon_i, \varepsilon_G + \beta_j^{G*} \varepsilon_j}(x, y) dx dy, \end{aligned}$$

where $f_{\varepsilon_G + \beta_i^{G*} \varepsilon_i, \varepsilon_G + \beta_j^{G*} \varepsilon_j}(x, y)$ is the density function of bivariate normal distribution of random variables $\varepsilon_G + \beta_i^{G*} \varepsilon_i$ and $\varepsilon_G + \beta_j^{G*} \varepsilon_j$. Finally,

$$\begin{aligned} q_i &= \rho_i + \sum_{j \neq i} \rho_j \left(1 - \frac{1}{\pi} \arctan \sqrt{(\beta_i^{G*})^2 + (\beta_j^{G*})^2 + (\beta_i^{G*} \beta_j^{G*})^2} \right) = \\ &1 - \frac{1}{\pi} \sum_{j \neq i} \rho_j \arctan \sqrt{(\beta_i^{G*})^2 + (\beta_j^{G*})^2 + (\beta_i^{G*} \beta_j^{G*})^2}, \end{aligned}$$

which completes the proof. ■

Region	mun.	years	y_{it}			ρ_i				x_{it}		
			obs.	mean	std.	obs.	mean	std.	max	obs.	mean	std.
					dev.							dev.
Aragon	731	87-07	1084	0.16	0.37	4386	0.001	0.019	0.50	4380	14.14	5.74
Asturias	78	83-07	334	0.11	0.31	546	0.013	0.037	0.25	546	10.45	5.39
Balearic Islands	67	95-07	95	0.20	0.40	268	0.015	0.046	0.38	268	18.78	3.92
Cantabria	102	87-03	140	0.19	0.39	510	0.010	0.035	0.34	510	10.99	4.29
Castile-La Mancha	919	87-07	2299	0.24	0.43	5514	0.001	0.004	0.08	5512	11.17	4.23
Extremadura	382	83-99	1019	0.10	0.30	1910	0.003	0.008	0.12	1909	5.76	2.50
La Rioja	174	87-03	390	0.19	0.39	870	0.006	0.037	0.48	870	14.06	6.54
Com. of Madrid	179	91-07	458	0.19	0.39	895	0.006	0.039	0.52	895	19.49	6.52
Region of Murcia	45	91-07	199	0.12	0.33	225	0.022	0.049	0.31	225	12.55	3.84
Valencian Com.	541	95-07	1291	0.21	0.41	2164	0.002	0.009	0.17	2148	15.71	3.82
Whole sample	3218		7309	0.19	0.39	17288	0.003	0.020	0.52	17263	12.57	5.82

Table 3: Sample characteristics and descriptive statistics.

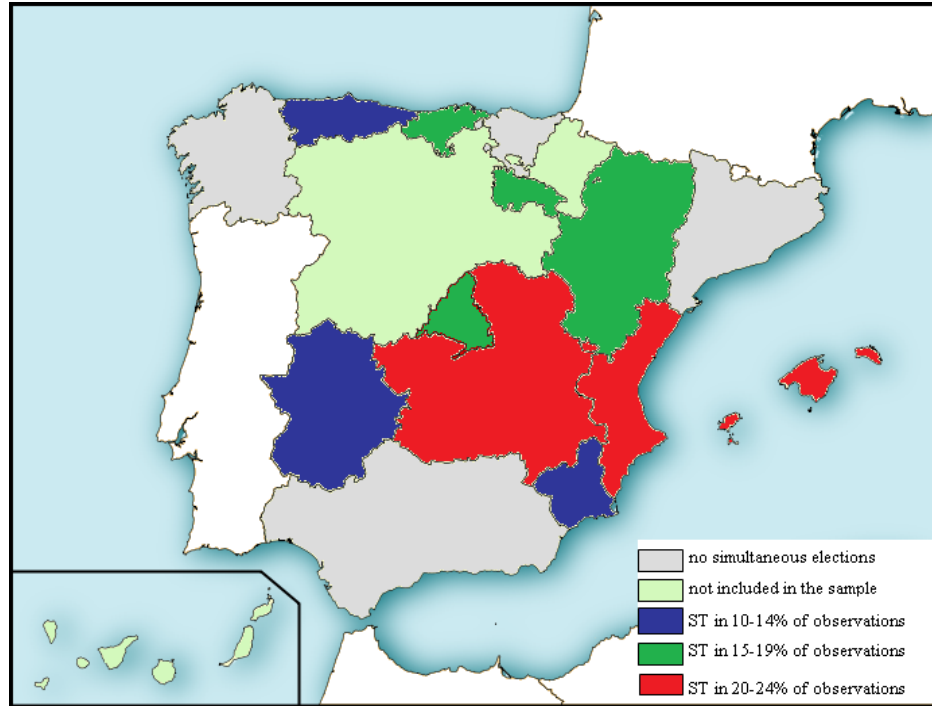


Figure 4: Percentage of split-ticket voting (ST) in the regions included in the sample.

CHAPTER II

SINCERE LOBBYING FORMATION

2.1 Introduction

The literature claims that wealthy citizens influence policy disproportionately because the willingness of the rich to make higher campaign contributions than the poor causes policy-makers to adopt positions the rich prefer (see, for example, Domhoff 1983 and Mills 1956, and, for the more recent contribution, Glazer and Gradstein 2005). However, in the ranking of the "Top 50" US political action committees (PACs) by contributions to candidates in 1999-2000 there are PACs listed whose members are more likely to be "poor" rather than "rich": Laborers' Political League, International Brotherhood of Electrical Workers Committee on Political Education, United Brotherhood of Carpenters and Joiners of America, etc.¹ Thus, the poor take an active part in lobbying to bias policy in their favor. Campante and Ferreira (2007) show that lobbying does favor the poor in the case where they have a comparative advantage in politics, rather than in production. However, I believe that the core of the problem is to model a decision rule that individuals use to decide whether to participate in lobbying or not. The aim of this paper is to build a tractable framework to explain this phenomenon analyzing individuals' decision to take part in lobbying activities.

The most prevalent formal literature approach builds on the assumption that lobbies influence political decisions through contributions (see Baron 1989, Becker 1983, 1985, Snyder 1990). The reviews of this and alternative approaches can be found in Austen-Smith (1997), Grossman and Helpman (2001), and Persson and Tabellini (2002). One wonders why some special interest groups have organized into lobbies and others have not. Olson (1971) identified some important issues of the problem. On the one hand, individuals with similar policy preferences can jointly influence policy outcome. On the other hand, there

¹Data from the Federal Election Committee. Available online at <http://www.fec.gov/press/press2001/053101pacfund/053101pacfund.html> (accessed January 23, 2008).

is always the strong temptation to free ride. Unfortunately, to solve a free-riding problem with a large number of individuals one needs to use coalition formation theory that proves to be very complicated. So the recent literature has addressed the question of lobbying formation in different contexts. Some authors focus mainly on formation of lobbies from exogenously given special interest groups (see Drazen et al. 2007, Felli and Merlo 2006, 2007, Laussel 2006 and Mitra 1999). Others address the problem of individuals' choice to lobby in some way. For example, Damania and Fredriksson (2000, 2003) and Magee (2002) analyze incentives for two firms and for n identical firms, respectively, to organize into a single industry lobby to affect policy outcomes. In turn, Bombardini (2008) proposes "Optimal Lobby Criterion" that reads: it is optimal for a firm to "join the lobby" if the joint surplus of a perspective member firm and the lobby is higher under firm participation. Glazer and Gradstein (2005) study the heterogeneous individuals' decision to make campaign contributions and show that people who contribute the most are extremists.

I develop a model of special interest politics analyzing individual's decision to participate in lobbying to influence public goods provision. I study complete information model with one-dimensional policy space. The incumbent government, which is either utilitarian or pro-median, cares about its utility and lobbies' contribution payments. Individuals are assumed to be heterogeneous in income, and I refer to low-income individuals as the poor and to high-income individuals as the rich. Accordingly, two lobbies can be organized: the lobby of the poor and the lobby of the rich. Moreover, I assume that there is no cost of forming lobbies and the lobbying mechanism is modeled as a common-agency problem à la Grossman and Helpman (1994). In equilibrium each individual either belongs to one of the two lobbies or does not participate in lobbying activities. I propose an intuitive condition for equilibrium termed *sincere lobbying formation*: an equilibrium occurs only if no lobby member would prefer her lobby to stop existing. This condition is obviously a necessary condition for equilibrium. Note that Alesina and Rosenthal's conditional sincerity condition for voter equilibrium applies a similar concept in voting context (see Alesina and Rosenthal 1995, 1996).

Why will individuals behave sincerely forming lobby groups rather than be involved in

free-riding? One possible explanation could be that individuals simply enjoy participating in special interest politics, unless they cannot afford it. In light of this fact, individuals may gain some personal satisfaction from showing allegiance to their special interest group. Another possible answer captures the idea of social norm individual behavior, that is, individuals take part in lobbying activities (unless they are better off without any lobby), because it is a social norm of the society. In other words, the social norm may advise: one should join a lobby if the gain one gets from lobbying activities is higher than the fee one is to pay as a lobby member. Alternatively, one can think of an ethical society where individuals bear a very high psychological cost if they engage in free-riding. So unless the gain from free-riding is considerably high, citizens will refrain from free-riding to avoid this psychological cost. In his turn, Smith (2000) in a systematical analysis of postwar lawmaking, shows that the public does overcome free-riding problem in the issues that affect the interests of the majority of the population such as tax rates, air pollution, and product liability.

I show that individuals with more extreme income levels are more likely to be involved in lobbying activities. To be more specific, in equilibrium each lobby is characterized by a threshold level of income such that all individuals with higher (for the lobby of the rich) or lower (for the lobby of the poor) income participate in lobbying activities. This is in line with the results of Glazer and Gradstein (2005) and McCarty et al. (2006) that extremists want to contribute the most.

I solve the model numerically, assuming that the gross income has lognormal distribution and using US gross income descriptive statistics from the Luxembourg Income Study. I find that the institute of lobbying does not necessarily favor wealthy citizens. On the contrary, in the case of a utilitarian government the final policy outcome favors the poor (in comparison with the socially optimal one). Accordingly, the lobby of the poor is more numerous and makes higher total campaign contributions than the lobby of the rich, while per member contribution is greater in the lobby of the rich. For example, for the United States the model predicts that the lobby of the poor is almost 4 times more numerous than the lobby of the rich, total contributions of the poor are 1.08 times higher than total contributions of

the rich, while each member of the lobby of the rich contributes 3.64 times more than each member of the lobby of the poor.

In the case of a pro-median government my results are in line with the existing literature: the final policy outcome does favor the rich (in comparison with the one preferred by the median voter). Although the lobby of the poor is more numerous, it contributes in total and per member less than the lobby of the rich. For example, for the United States the numerical results read: the lobby of the poor is more than twice the size of the lobby of the rich, while total contributions of the rich are 3.56 times higher than the ones of the poor, with each member of the lobby of the rich contributing almost 8 times more than each member of the lobby of the poor.

So the conceptual difference in my results comes from the assumptions on the government preferences: under a utilitarian government lobbying favors the poor, while under a pro-median government lobbying favors the rich. The reason for this is quite intuitive and due to the fact that income distribution is recognized to be skewed to the right. Pro-median government tends to satisfy the median voter preferences, while utilitarian government implements the policy preferred by the individual with mean income. Therefore, under a pro-median government, the rich have more stake in the policy and coordinate better, since without lobbying the policy outcome is not in their favor. Accordingly, under a utilitarian government, the poor have more stake in the policy and coordinate better, since the mean individual preferred policy does not favor them. Still, with or without lobbying, the poor would prefer a pro-median government to a utilitarian one, while the rich would prefer a utilitarian government to a pro-median one.

If the government cares only about contribution payments all individuals participate in lobbying: individuals with income lower than the mean one belong to the lobby of the poor, and individuals with income higher than the mean one belong to the lobby of the rich. This happens because individuals know that the only way to get favorable policy outcome is lobbying: the government does not care about citizens' wellbeing at all. In this case political competition results in socially optimal outcome.

I analyze how the degree of income inequality in a society affects the composition of

lobbies and the final outcome in the equilibrium. The model predicts that the less egalitarian the society is, the more (resp. the less) numerous the lobby of the poor (resp. the rich) is, and the higher the final tax rate for both utilitarian and pro-median government. Still the qualitative results stay the same: in the case of a utilitarian government, lobbying favors the poor while in the case of a pro-median government lobbying favors the rich.

Now turn to the assumptions of the model. I assume that individuals' utility is quasilinear in income. I want to concentrate on relative rather than absolute magnitudes of lobbying formation, so this assumption allows to isolate the effects of interest (for example, how lobbying formation depends on the shape of income distribution). Would my results change with concave utility function? The answer to this question is not obvious. In this case the rich would value each additional unit of income less than the poor so one expects that the rich will contribute more than the poor. However, in the case where the government is utilitarian, it cares about the social welfare of all the individuals, and the poor would value a slight increase in taxes much more than the rich would value a slight (same) decrease in taxes, so it would be much cheaper for the poor to buy influence than for the rich. What effect would dominate depends on the particular utility function. Still, I expect that my results will prevail for utility functions that are not too concave in income.

Summing up, I make the following contributions. First, I introduce a new condition for lobbying formation equilibrium, namely, *sincere lobbying formation* that reads: an equilibrium occurs if no lobby member would prefer her lobby to cease to exist. Second, I show that the institute of lobbying can favor the poor if office-holders do not care about their reelection prospects and do care about the citizens' welfare (the case of a utilitarian government). However, if the elections are coming and policymakers want to be reelected for the next term, the wealthy citizens get more political power and the final policy outcome favors the rich (the case of a pro-median government). Finally, I find that political competition can result in socially optimal outcome. It happens if office-holders care only about contribution payments.

The rest of the paper is organized as follows. Section 2.2 lays out a simple model of public goods provision. Section 2.3 describes the common-agency model of lobbying. Section 2.4

develops the sincere lobbying formation concept. Section 2.5 contains numerical solutions for lognormal distribution of gross income. Finally, Section 2.6 concludes.

2.2 *A Simple Model of Public Goods Provision*

Consider a society inhabited by a large number (formally a continuum) of individuals, where I normalize the size (mass) of the population to unity. Individuals differ in their income x . I assume that x is distributed in the population according to a smooth, at least twice differentiable cumulative distribution function $F(\cdot)$ with mean \hat{x} and support $[0, \infty)$. The corresponding density function is denoted by $f(\cdot)$. Assume further that x is skewed to the right, in accordance with evidence from virtually every country.

Each individual with income x has the same quasi-linear preferences over private consumption $c(x)$ and publicly provided goods g , which is given by

$$u(c, g) = c + \sqrt{g}.$$

One can interpret g in different ways, as publicly provided private goods, or traditional public goods. Let g measure spending per capita. Government spending is financed by taxing the income of every individual at a common rate $t \in [0, 1]$. Then consumption differs according to

$$c(x) = (1 - t)x,$$

and the government budget constraint is then simply

$$t\hat{x} = g.$$

Then the tax preferences of individual x read

$$u(x, t) = (1 - t)x + \sqrt{t\hat{x}},$$

that are concave in tax, implying that every individual x has a uniquely preferred tax rate t_x :

$$t_x = \begin{cases} 1 & \text{if } x \in [0, \frac{1}{2}\sqrt{\hat{x}}] \\ \frac{\hat{x}}{4x^2} & \text{if } x \in (\frac{1}{2}\sqrt{\hat{x}}, \infty). \end{cases}$$

Richer individuals want a smaller government because, with taxes proportional to income, they pay a larger share of the tax burden.²

Let us formulate a normative benchmark. As a basis for this benchmark, consider a utilitarian social welfare function that simply integrates over the welfare of all individuals:

$$U^o(t) = \int_0^\infty u(x, t) f(x) dx = (1 - t) \hat{x} + \sqrt{t\hat{x}},$$

where the last term is just the utility of the individual with mean income. Then the socially optimal tax rate coincides with the tax desired by the mean individual $t^o = \frac{1}{4\hat{x}}$.

Alternatively, the median voter preferences can be considered as a benchmark:

$$U^m(t) = (1 - t) x_m + \sqrt{t\hat{x}},$$

where x_m stands for the median voter income and $F(x_m) = \frac{1}{2}$. Then the median voter preferred tax rate reads $t^m = \frac{\hat{x}}{4x_m^2}$ (that is assumed to be less than 1).

In what follows I work with these two alternative benchmarks.

2.3 Lobby Groups

I focus on lobbying activities in the context of the common-agency model of Bernheim and Whinston (1986) adapted to lobbying by Grossman and Helpman (1994). In this approach, lobbying is modeled as "menu auction" where lobbies confront a government with contribution schedules that map any possible policy into a contribution payment. Several authors have applied the common-agency model of lobbying to study trade policy, commodity taxation, provision of local public goods and other policies (see Dixit et al. 1997, Grossman and Helpman 1996, Helpman and Persson 2001, Persson 1998). I use this approach as well.

I assume that just two lobbies can be formed: a lobby of low-income individuals (the poor) given by set P , and a lobby of high-income individuals (the rich) given by set R . Denote by L the set of organized lobbies. In this section I leave aside lobbying formation considerations and assume that each individual can join one lobby using a decision rule to

²I assume that individuals' utility is quasilinear in income to isolate the effects of income inequality on lobbying formation.

be specified below. Suppose further that the lobbies care about the sum of their members' welfare. Thus, the gross objective function of each lobby $l \in L$ is given by:

$$U^l(t) = \int_{x \in l} u(x, t) f(x) dx.$$

At the first stage of the game each lobby $l \in L$, non-cooperatively and simultaneously, presents its common agent, the government, with a contribution schedule $C^l(t)$ giving a binding promise of payment conditional on the chosen tax rate. Following the literature, I concentrate on (globally) truthful contribution schedules that satisfy:

$$C^l(t) = \max \left[U^l(t) - b^l, 0 \right],$$

where b^l is a constant chosen optimally by lobby l . The objective of lobby l is to maximize the net welfare of its members, namely $U^l(t) - C^l(t)$.

At the second stage, the government sets t to maximize a weighted sum of its utility and contributions:

$$\alpha U(t) + \sum_{l \in L} C^l(t), \quad \alpha \geq 0. \quad (3)$$

An equilibrium of the game is a Subgame perfect Nash equilibrium in the contribution schedules and the chosen tax rate.

I analyze two alternative scenarios, namely an utilitarian government, $U(t) = U^o(t)$, and a pro-median government, $U(t) = U^m(t)$. In the former case, the government cares about the individuals' welfare and about contribution payments. In the latter case, the government is concerned both about total amount of contributions and about its chances of being reelected. In political economy literature the government concerned about its reelection prospects maximizes the probability of winning the election. However, in this framework the election itself is not modeled, so it is more convenient if the government's objective function gives greater weight to individuals that are believed to determine the election outcome. Then the closer the government's objective is to the median-voter preferred policy, the higher the probability to win the election. Alternatively, one can think of the situation where office-holders are "citizen-candidates" (as in Besley and Coate 1997 or Osborne and Slivinski 1996) who share the preferences of either an individual with mean income or the median voter.

To derive an equilibrium in truthful strategies, I use the fact that equilibrium tax rate is Pareto optimal in the bilateral relation between the government and each lobby.³ Therefore, the equilibrium tax rate t maximizes the sum of the organized lobbies' net welfare $\sum_{l \in L} (U^l(t) - C^l(t))$ and the government objective (3). Then the optimal tax rate maximizes:

$$\alpha U(t) + \sum_{l \in L} U^l(t). \quad (4)$$

The first-order condition of (4) yields the equilibrium tax rate t^* :

$$t^* = \arg \max_{t \in [0,1]} \left(\alpha U(t) + \sum_{l \in L} U^l(t) \right).$$

To find the contribution levels in the equilibrium, define by t^{-l} the tax rate that would emerge if the contribution offered by lobby l were zero, so

$$t^{-l} = \arg \max_{t \in [0,1]} \left(\alpha U(t) + \sum_{\substack{i \in L \\ i \neq l}} C^i(t) \right).$$

In other words, t^{-l} is the tax rate that would emerge if lobby l were not formed.

Lobby l will raise its b^l to the point where the government is just indifferent between choosing the tax rate t^{-l} and choosing the equilibrium tax rate t^* , that is

$$\alpha U(t^{-l}) + \sum_{\substack{i \in L \\ i \neq l}} C^i(t^{-l}) = \alpha U(t^*) + \sum_{i \in L} C^i(t^*) \text{ for all } l \in L.$$

These two sets of equations allow us to solve for the lobbies' contributions in the equilibrium:

$$\begin{aligned} C^{P*} &= \alpha (U(t^{-P}) - U(t^*)) + U^R(t^{-P}) - U^R(t^*) \\ C^{R*} &= \alpha (U(t^{-R}) - U(t^*)) + U^P(t^{-R}) - U^P(t^*). \end{aligned} \quad (5)$$

In the case where there is just one organized lobby, the government derives exactly the same utility as it would have achieved without any contribution. Thus, a lobby that faces no competition captures the entire surplus from lobbying activities. If all individuals participate in lobbying, the government captures the entire surplus from lobbying activities and each lobby pays according to the political strength of its rival.

³See Bernheim and Whinston (1986) and Dixit et al. (1997) for the proof.

So far my analysis leaves aside the crucial issue of lobbying formation. I study this question in the following section.

2.4 *Sincere Lobbying Formation*

I assume that lobby P 's goal is to defend special interests of the poor while lobby R aims to defend special interests of the rich. In my simple model the special interests vary with preferences over tax rate, that is, in general the poor prefer bigger government and, thus, a higher tax rate, while the rich want smaller government and a lower tax rate. To reflect this conceptual difference between the two lobbies, I assume that in equilibrium lobby P and lobby R make contributions to the government in order to raise and to drop, respectively, the final tax rate, that is, $t^{-P} \leq t^* \leq t^{-R}$.

In this game the choice of each individual is either to be a member of one of two lobbies, P or R , or not to participate in lobbying activities at all. I assume that each individual can belong just to one lobby group since in my model lobbies represent opposite interests. There is no fixed cost of forming lobbies. If an individual belongs to a lobby, her welfare is taken into account when the lobby develops a contribution schedule, but she should bear a contribution burden which is the same for all the lobby members.⁴ How do individuals manage to solve the coordination problem while making their choice? I assume that the coordination has a simple form that I call *sincere lobbying formation*.⁵

Sincere Lobbying Formation Condition: An equilibrium occurs only if no lobby member would prefer her lobby to cease to exist.

It is evident that the condition should hold in equilibrium: if a lobby member would like her lobby to cease to exist, then she is "lobbying" in the "wrong" way given her expectations

⁴In this quasilinear model it is reasonable to assume that contributions should be proportional to a marginal utility of income that is the same for all individuals.

⁵Alternatively, one can think of a society inhabited by individuals of two types with the size of each group normalized to unity. The first type individuals are free-riders: they never participate in lobbying activities. The second type individuals are faithful to their special interest group: they join a lobby if the gain they get from lobbying activities exceeds the contribution fee. The analysis stays the same for this alternative interpretation of the model.

and preferences. Formally, the sincere lobbying formation condition reads:

$$\begin{aligned} \text{if } x \text{ belongs to lobby } P, \text{ then } u(x, t^*) - \frac{C^{P*}}{\int_{z \in P} f(z) dz} &> u(x, t^{-P}) \\ \text{if } x \text{ belongs to lobby } R, \text{ then } u(x, t^*) - \frac{C^{R*}}{\int_{z \in R} f(z) dz} &> u(x, t^{-R}). \end{aligned}$$

Term indifferent individuals as sincere indifferent poor, π , and sincere indifferent rich, ρ , such that

$$\begin{aligned} u(\pi, t^*) - \frac{C^{P*}}{\int_{z \in P} f(z) dz} &= u(\pi, t^{-P}) \\ u(\rho, t^*) - \frac{C^{R*}}{\int_{z \in R} f(z) dz} &= u(\rho, t^{-R}). \end{aligned} \tag{6}$$

Given this equilibrium concept there can be multiple equilibria. In what follows I consider the largest possible lobbies, that is, if there are more than one π and more than one ρ satisfying conditions (6), then I call sincere indifferent poor the highest π and sincere indifferent rich the lowest ρ .

Then I establish:

Lemma 7 *If in equilibrium there exist lobby P and lobby R then $P = \{x | x \in [0, \pi)\}$ and $R = \{x | x \in (\rho, \infty)\}$.*

The formal proof can be found in the Appendix.

Solve the game backwards. Suppose that in equilibrium there exist lobby $P = \{x | x \in [0, \pi)\}$ and lobby $R = \{x | x \in (\rho, \infty)\}$. To reflect the fact that in equilibrium each individual can belong either to one lobby or to no lobby I assume that $\pi \leq \rho$. The final goal is to find these π and ρ .

Now I introduce new pieces of notation. I denote by $s^l(\cdot)$ the size of lobby l . Then $s^P(\pi) \equiv F(\pi)$ and $s^R(\rho) \equiv 1 - F(\rho)$. The aggregate income in lobby P (resp. R) is $W^P(\pi) \equiv \int_0^\pi x f(x) dx$ (resp. $W^R(\rho) \equiv \int_\rho^\infty x f(x) dx$).

The lobbies gross objective functions read:

$$\begin{aligned} U^P(t, \pi) &= \int_0^\pi u(x, t) f(x) dx = (1-t) W^P(\pi) + \sqrt{t\hat{x}} s^P(\pi) \\ U^R(t, \rho) &= \int_\rho^\infty u(x, t) f(x) dx = (1-t) W^R(\rho) + \sqrt{t\hat{x}} s^R(\rho). \end{aligned}$$

Then the lobbies develop truthful contribution schedules to offer to the government that chooses a final tax rate t^* :

$$t^*(\pi, \rho) \equiv \arg \max_{t \in [0,1]} (\alpha U(t) + U^P(t, \pi) + U^R(t, \rho)) = \frac{\hat{x}}{4} \left(\frac{\alpha + s^P(\pi) + s^R(\rho)}{\alpha\chi + W^P(\pi) + W^R(\rho)} \right)^2 \quad (7)$$

where

$$\chi = \begin{cases} \hat{x} & \text{if } U(t) = U^o(t) \\ x_m & \text{if } U(t) = U^m(t). \end{cases}$$

If lobby l ($l = P, R$) were not around the tax rate $t^{-l}(\cdot)$ would emerge:

$$t^{-P}(\rho) \equiv \arg \max_{t \in [0,1]} (\alpha U(t) + U^R(t, \rho)) = t^*(0, \rho) = \frac{\hat{x}}{4} \left(\frac{\alpha + s^R(\rho)}{\alpha\chi + W^R(\rho)} \right)^2 \quad (8)$$

$$t^{-R}(\pi) \equiv \arg \max_{t \in [0,1]} (\alpha U(t) + U^P(t, \pi)) = t^*(\pi, \infty) = \frac{\hat{x}}{4} \left(\frac{\alpha + s^P(\pi)}{\alpha\chi + W^P(\pi)} \right)^2. \quad (9)$$

Denote by τ the tax rate that would emerge if there were no lobbies formed. It depends on the government type ($\tau = t^o$ or $\tau = t^m$) and reads $\tau = \frac{\hat{x}}{4\chi^2}$.

In the case of the utilitarian government, $U(t) = U^o(t)$, the equilibrium tax rate can be efficient: $t^*|_{\chi=\hat{x}} = t^o$. This happens when all individuals participate in lobbying. In this case lobbies "neutralize" one another, so that R 's bids for a smaller government are matched in the equilibrium by P 's bids for a bigger government, and political competition results in socially optimal outcome. Nonetheless, each lobby must make a positive contribution in order to induce the government to choose this outcome rather than one that would be still worse from its perspective. If just one lobby were organized, the equilibrium tax rate would differ from the social optimum in favor of the organized group.

In general, the following inequalities hold:

$$\begin{aligned} t^{-P}(\rho) &< t^*(\pi, \rho) < t^{-R}(\pi) \\ t^{-P}(\rho) &< \tau < t^{-R}(\pi), \end{aligned}$$

while the relationship between $t^*(\pi, \rho)$ and τ is as follows:

$$t^*(\pi, \rho) \begin{matrix} \geq \\ \leq \end{matrix} \tau \text{ iff } \chi \begin{matrix} \geq \\ \leq \end{matrix} \frac{W^P(\pi) + W^R(\rho)}{s^P(\pi) + s^R(\rho)}. \quad (10)$$

Condition (10) reads: Lobbying favors the poor (resp. the rich), in other words, the final tax rate is higher (resp. lower) than the tax rate that would emerge if there were no lobbies

formed, if and only if the mean income in the society (in the case of utilitarian government) or the median-voter income (in the case of pro-median government) is higher (resp. lower) than the mean income in both lobbies. Lobbying does not affect the final tax rate if and only if the mean income in the society (in the case of utilitarian government) or the median-voter income (in the case of pro-median government) is equal to the mean income in both lobbies. Thus, the final tax rate goes in favor of a lobby with higher relative political strength.

I use (5) to find the lobbies' contributions in the equilibrium:

$$C^{P*}(\pi, \rho) = \frac{\hat{x}}{4(\alpha\chi + W^R(\rho))} \cdot \left(\frac{(\alpha + s^R(\rho))(\alpha\chi + W^P(\pi) + W^R(\rho)) - (\alpha + s^P(\pi) + s^R(\rho))(\alpha\chi + W^R(\rho))}{\alpha\chi + W^P(\pi) + W^R(\rho)} \right)^2$$

$$C^{R*}(\pi, \rho) = \frac{\hat{x}}{4(\alpha\chi + W^P(\pi))} \cdot \left(\frac{(\alpha + s^P(\pi))(\alpha\chi + W^P(\pi) + W^R(\rho)) - (\alpha + s^P(\pi) + s^R(\rho))(\alpha\chi + W^P(\pi))}{\alpha\chi + W^P(\pi) + W^R(\rho)} \right)^2.$$

Given the results above I turn now to the lobbying formation stage of the game. Formally, for sincere indifferent poor π and sincere indifferent rich ρ , the following two conditions must hold:

$$u(\pi, t^*(\pi, \rho)) - \frac{C^{P*}(\pi, \rho)}{s^P(\pi)} = u(\pi, t^{-P}(\rho))$$

$$u(\rho, t^*(\pi, \rho)) - \frac{C^{R*}(\pi, \rho)}{s^R(\rho)} = u(\rho, t^{-R}(\pi)),$$

that yield the system of two equations with two unknowns π and ρ :

$$\sqrt{\frac{t^{-P}(\rho)}{\hat{x}}} + \sqrt{\frac{t^*(\pi, \rho)}{\hat{x}}} = \frac{\alpha + s^P(\pi) + s^R(\rho)}{\pi s^P(\pi) + \alpha\chi + W^R(\rho)}$$

$$\sqrt{\frac{t^{-R}(\pi)}{\hat{x}}} + \sqrt{\frac{t^*(\pi, \rho)}{\hat{x}}} = \frac{\alpha + s^P(\pi) + s^R(\rho)}{\rho s^R(\rho) + \alpha\chi + W^P(\pi)}.$$

After plugging in the expressions for $t^*(\pi, \rho)$, $t^{-P}(\rho)$, $t^{-R}(\pi)$ from (7), (8) and (9) this system reads:

$$\frac{\alpha + s^R(\rho)}{\alpha + s^P(\pi) + s^R(\rho)} \frac{\alpha\chi + W^P(\pi) + W^R(\rho)}{\alpha\chi + W^R(\rho)} = \frac{\alpha\chi + W^R(\rho) + 2W^P(\pi) - \pi s^P(\pi)}{\alpha\chi + W^R(\rho) + \pi s^P(\pi)}$$

$$\frac{\alpha + s^P(\pi)}{\alpha + s^P(\pi) + s^R(\rho)} \frac{\alpha\chi + W^P(\pi) + W^R(\rho)}{\alpha\chi + W^P(\pi)} = \frac{\alpha\chi + W^P(\pi) + 2W^R(\rho) - \rho s^R(\rho)}{\alpha\chi + W^P(\pi) + \rho s^R(\rho)}.$$

In general, it is not straightforward to find an explicit form solution for this system of two equations with two unknowns. In what follows, I assume the lognormal distribution of income and solve for π and ρ numerically.

2.5 Numerical Solution: Lognormal Distribution of Income

The lognormal distribution is very popular in modeling applications, when the variable of interest is skewed to the right. I use this distribution as well. Formally, I assume that x has a lognormal distribution (that is, $\ln x \sim N(\mu, \sigma^2)$). Then the density function reads

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad 0 < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$$

with mean

$$\hat{x} = e^{\mu + \frac{\sigma^2}{2}}$$

and variance

$$Var = e^{2(\mu + \sigma^2)} - e^{2\mu + \sigma^2}.$$

To generate the distribution I use gross income descriptive statistics (in particular, mean and standard deviation) for the United States from the Luxembourg Income Study (LIS) dataset for households.⁶ In LIS dataset gross income amounts are in national currency units on the year of survey. Since my primary goal is to see the relative magnitudes of lobbying formation, I normalize gross income statistics for the ease of presentation (see Table 4 in the Appendix for the original LIS descriptive statistics, normalized descriptive statistics (\hat{x} , \sqrt{Var} , x_m), μ and σ for the normalized descriptive statistics, and socially optimal tax rate t^o and median voter preferred tax rate t^m).

First, I consider a utilitarian government, $\chi = \hat{x}$. Columns 1 and 2 in Table 5 in the Appendix contain numerical results and relative magnitudes of lobbying formation for an utilitarian government with $\alpha = 1$ and $\alpha = 100$. Note that the model predicts that in this case the final tax rate, t^* , favors the lobby of the poor P , that is $t^* > t^o$. Accordingly, the total contribution of lobby P exceeds the total contribution of lobby R , while per member contribution is higher in lobby R than in lobby P . If $\alpha = 1$ the model predicts that around

⁶ Available online at <http://www.lisproject.org/techdoc.htm> (accessed December 11, 2007).

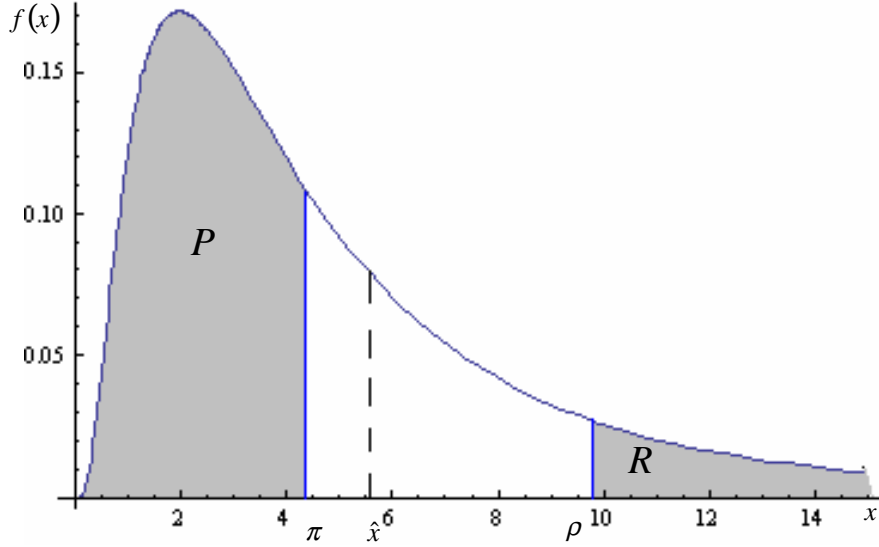


Figure 5: Sincere Lobbying Formation for US data: the case of utilitarian government with $\alpha = 1$.

half of the population joins the lobby of the poor P . As for the lobby of the rich R , it is much smaller: just around 14 % of the population. The lobby of the poor P is around 4 times bigger than the lobby of the rich R . As for contribution, the lobby of the poor P pays in total slightly more than the lobby of the rich R , while per member contribution is around 3.6 times higher for lobby R members than for lobby P members. Figure 5 pictures lobbying formation for US data in the case of utilitarian government with $\alpha = 1$.

When $\alpha = 100$, the government cares much more about social welfare than about contribution payments, so it is not easy for lobbies to buy influence. As a result, lobbies are smaller in size and contribute much less than in the case of $\alpha = 1$. The equilibrium tax rate is just slightly higher than the social optimal one and lobby P in total contributes twice as much as lobby R .

The literature claims that the institute of lobbying does favor richer strata of the society (see Domhoff 1983 and Mills 1956). However, my results indicate that it is not necessarily the case for the utilitarian government. My model of sincere lobbying formation predicts that lobbying favors poorer individuals, that is, the final outcome of lobbying formation is in favor of the lobby of the poor in comparison with the socially optimal one. Moreover, the

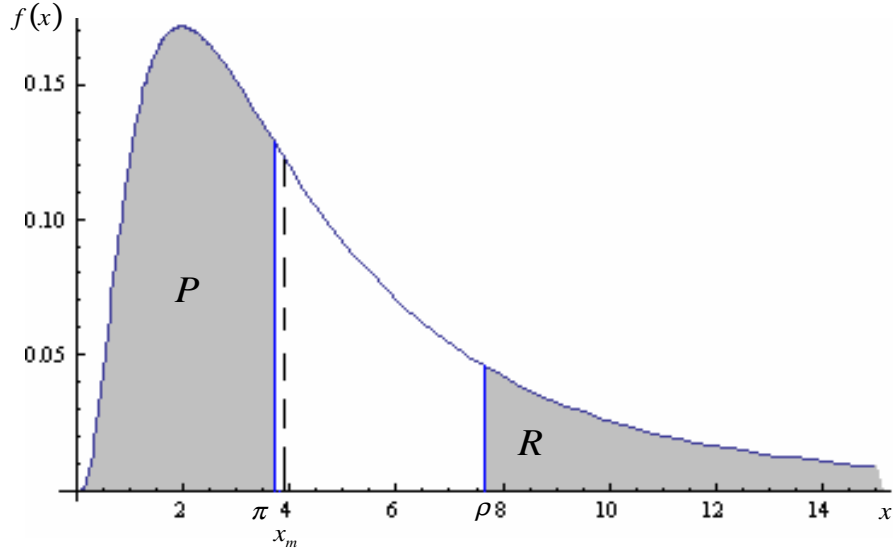


Figure 6: Sincere Lobbying Formation for US data: the case of pro-median government with $\alpha = 1$.

lobby of the poor is considerably bigger and contributes more than the lobby of the rich, while per member contribution is higher in the lobby of the rich.

Now consider a pro-median government, $\chi = x_m$. Columns 3 and 4 in Table 5 in the Appendix present the results for a pro-median government with $\alpha = 1$ and $\alpha = 100$. The results indicate that in this case lobbying does favor richer individuals, $t^* < t^m$. Lobby P is smaller in size and lobby R is bigger in size than in the case of the utilitarian government for corresponding values of α . Total contributions and per member contribution are higher in lobby R than in lobby P . If $\alpha = 1$ lobby P is twice bigger than lobby R . However, the lobby of the rich contributes around 3.5 times more in total and around 7.8 times more per member, than the lobby of the poor. Figure 6 pictures lobbying formation for the case of pro-median government with $\alpha = 1$ for US data.

When $\alpha = 100$, it is very difficult for lobbies to buy influence. Therefore, the lobbies are smaller in size and pay lower contributions than in the case of the pro-median government with $\alpha = 1$. Still, lobbying favors richer individuals since the equilibrium tax rate is slightly lower than the one preferred by the median voter, and the lobby of the rich pays 4.3 times higher total contribution than the lobby of the poor.

The results for the pro-median government are in line with the existing literature: lobbying does favor wealthy strata of the society. The final policy outcome is in favor of the lobby of the rich (in comparison with the one preferred by the median voter). The lobby of the poor is more numerous than the lobby of the rich. However, total contributions and per member contribution are higher in the lobby of the rich.

It is of interest to see the equilibrium evolution for $\alpha = 0$. Column 5 in Table 5 in the Appendix contains numerical results and relative magnitudes of lobbying formation for $\alpha = 0$. Here the government cares only about contribution payments and individuals know this. As a result in the equilibrium all individuals participate in lobbying: individuals with income lower than the mean one belong to the lobby of the poor, and individuals with income higher than the mean one belong to the lobby of the rich. This happens because individuals know that the only way to get favorable policy outcome is lobbying: the government does not care about citizens' wellbeing at all. Political competition results in socially optimal outcome, $t^* = t^o$.⁷ As I mentioned above, in this case in equilibrium lobbies "neutralize" one another. Both lobbies pay higher total contributions and higher per member contributions than in the case of $\alpha = 1$. However, the lobby of the rich, R , contributes in total more than the lobby of the poor, P , and lobby P to lobby R size ratio is equal to lobby R to lobby P total contribution ratio. This is because in equilibrium each lobby pays according to the political strength of its rival: each lobby must contribute an amount equal to the difference between what its rival could achieve without competition and what it actually achieves in equilibrium. Figures 7 and 8 in the Appendix depict the evolution of lobbying formation equilibrium with the change of $\alpha \geq 0$ for US data. Note that the final tax rate under pro-median government is higher than the one under utilitarian government with lobbying or without it. Therefore, the poor would prefer the pro-median government to the utilitarian one in spite of the fact that under the latter they could influence final policy in their favor by lobbying. In their turn, the rich would prefer the utilitarian government to the pro-median one, even when they could lobby more successfully under a pro-median

⁷The result that policy outcome is socially optimal when all individuals participate in lobbying is due to Grossman and Helpman (1994). My contribution here is to specify that this happens under a "corrupted" government that cares only about lobbies' donations.

government.

How does the degree of inequality affect sizes of lobbies and policy outcome in the equilibrium? Figures 9 and 10 in the Appendix represent, respectively, the evolution of lobbies' sizes and tax rate in the equilibrium with the change of standard deviation given the mean.⁸ The less egalitarian the society (the higher the standard deviation given the mean), the more numerous the lobby of the poor is and the less numerous the lobby of the rich is. This is due to the fact that there are more poor and fewer rich individuals in less egalitarian societies. As for the equilibrium tax rate, it is increasing in the degree of inequality both for utilitarian and pro-median governments. This happens just for the same reason: there are more poor individuals, the lobby of the poor is more numerous so it can influence final policy outcome more successfully. Note that the higher the income inequality is in the society, the more political influence the poor have under a utilitarian government.

2.6 Conclusion

The paper studies the impact of lobbying on a public goods provision. I propose a new equilibrium condition for lobbying formation, namely, *sincere lobbying formation*: an equilibrium occurs only if no lobby member would prefer her lobby to stop existing. Lobbying is modeled as a common-agency problem, only two lobbies can be organized, and there is no cost of forming lobbies. The model predicts that individuals with more extreme preferences are more likely to participate in lobbying. I solve the model numerically with the US data from the Luxembourg Income Study to show that lobbying does not necessarily favor wealthy citizens. The results indicate that if policymakers do not care about their reelection prospects and do care about the individuals' welfare (utilitarian government), the final policy outcome favors the poor (in comparison with the socially optimal one). In this case the lobby of the poor is bigger in size and makes higher total campaign contributions than the lobby of the rich, while per member contribution is greater in the lobby of the rich. However, if the elections are close and policymakers want to be reelected (pro-median

⁸The mean comes from US data: $\hat{x} = 5.611469$.

government), the final policy outcome does favor the rich (in comparison with the one preferred by the median voter), which is in line with the existing literature. In spite of the fact that the lobby of the poor is more numerous, its total and per member contributions are lower than ones of the lobby of the rich. However, lobbying does not change final policy drastically: with or without lobbying the poor would prefer a pro-median government to a utilitarian one, while the rich would prefer a utilitarian government to a pro-median one. In the case where the government cares only about lobbies' contribution payments, all individuals participate in lobbying: individuals with income lower (resp. higher) than the mean one belong to the lobby of the poor (resp. the lobby of the rich). In this case political competition results in socially optimal outcome.

The degree of income inequality in the economy does affect the composition of lobbies and the final policy outcome in quantitative terms, namely, the less egalitarian the society, the more numerous the lobby of the poor, the less numerous the lobby of the rich, and the higher the final tax rate both for utilitarian and pro-median governments.

2.7 Appendix

Proof of Lemma 7. Assume that in equilibrium there exists lobby P . Then

$$P = \left\{ x \mid u(x, t^*) - \frac{C^{P*}}{\int_{z \in P} f(z) dz} > u(x, t^{-P}) \right\}.$$

After straightforward calculations and taking into account that $t^* \geq t^{-P}$, the last inequality reads

$$x < \frac{\sqrt{x}}{\sqrt{t^*} + \sqrt{t^{-P}}} - \frac{1}{t^* - t^{-P}} \frac{C^{P*}}{\int_{z \in P} f(z) dz} \equiv \pi.$$

Thus, in equilibrium lobby P satisfies $P = \{x \mid x \in [0, \pi)\}$.

If in equilibrium there exists lobby R , then

$$R = \left\{ x \mid u(x, t^*) - \frac{C^{R*}}{\int_{z \in R} f(z) dz} > u(x, t^{-R}) \right\}.$$

Taking into account that $t^* \leq t^{-R}$, the last inequality yields

$$x > \frac{\sqrt{x}}{\sqrt{t^*} + \sqrt{t^{-R}}} - \frac{1}{t^* - t^{-R}} \frac{C^{R*}}{\int_{z \in R} f(z) dz} \equiv \rho.$$

So, in equilibrium lobby R satisfies $R = \{x \mid x \in (\rho, \infty)\}$. ■

LIS Mean	LIS Std. Dev.	\hat{x}	\sqrt{Var}	x_m	μ	σ	t^o	t^m
56114.69	56463.514	5.611469	5.6463514	3.95559	1.37513	0.83628	0.0445516	0.0896588

Table 4: The US gross income descriptives from LIS household files (year 2000), normalized descriptives, corresponding parameters for lognormal distribution, socially optimal tax rate and median voter preferred tax rate.

α	utilitarian government		pro-median government		
	1	100	1	100	0
π	4.41232	3.93564	3.75018	2.87648	5.611469
s^P	0.551978	0.497587	0.474578	0.351628	0.662078
ρ	9.75917	12.7233	7.63875	10.5252	5.611469
s^R	0.140099	0.0812024	0.215658	0.120954	0.337922
t^*	0.0471963	0.0446328	0.0641447	0.0892857	0.0445516
C^{P*}	0.0165386	$2.20784 \cdot 10^{-4}$	0.0131437	$1.31976 \cdot 10^{-4}$	0.0396768
$\frac{C^{P*}}{s^P}$	0.0299624	$4.43709 \cdot 10^{-4}$	0.0276956	$3.75329 \cdot 10^{-4}$	0.0599277
C^{R*}	0.0152982	$10.6438 \cdot 10^{-5}$	0.0467925	$5.71429 \cdot 10^{-4}$	0.0777371
$\frac{C^{R*}}{s^R}$	0.109195	$13.1077 \cdot 10^{-4}$	0.216976	$4.72435 \cdot 10^{-3}$	0.230044
$\frac{s^P}{s^R}$	3.93991	6.12774	2.20061	2.90712	1.95926
$\frac{C^{P*}}{C^{R*}}$	1.08108	2.0743	-	-	-
$\frac{C^{R*}}{C^{P*}}$	-	-	3.56007	4.3298	1.95926
$\frac{\frac{C^{R*}}{s^R}}{\frac{C^{P*}}{s^P}}$	3.64442	2.95412	7.83431	12.5872	3.83869

Table 5: Numerical results and relative magnitudes of lobbying formation for utilitarian and pro-median governments with $\alpha = 1$, $\alpha = 100$ and $\alpha = 0$.

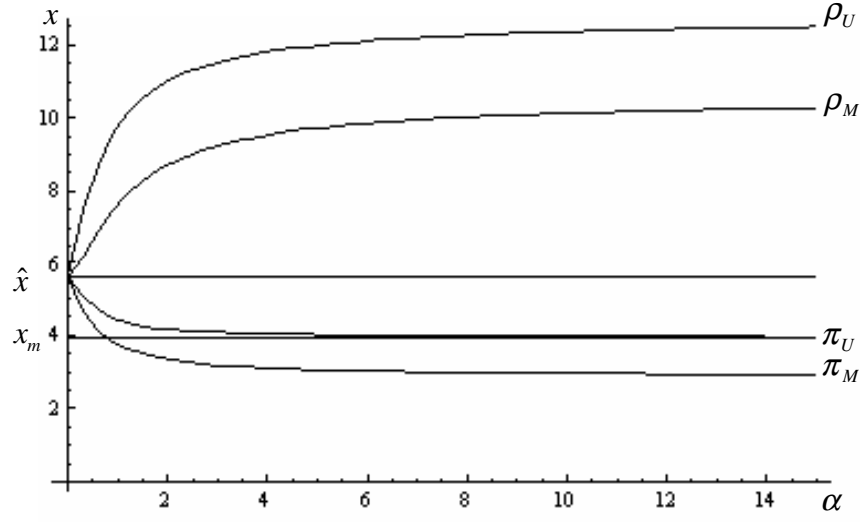


Figure 7: Sincere indifferent poor π and rich ρ as a function of α for US data: π_U, ρ_U (utilitarian government) and π_M, ρ_M (pro-median government).

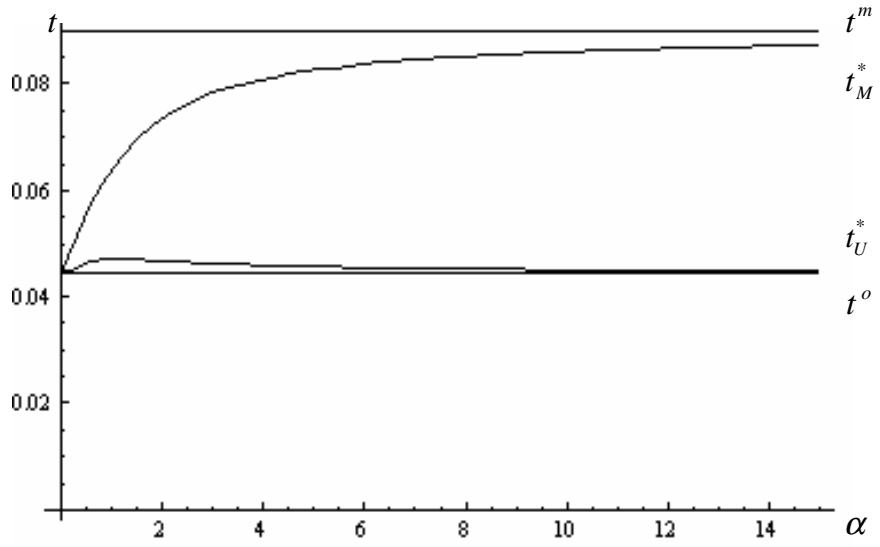


Figure 8: Equilibrium tax rate as a function of α for US data: t_U^* (utilitarian government) and t_M^* (pro-median government).

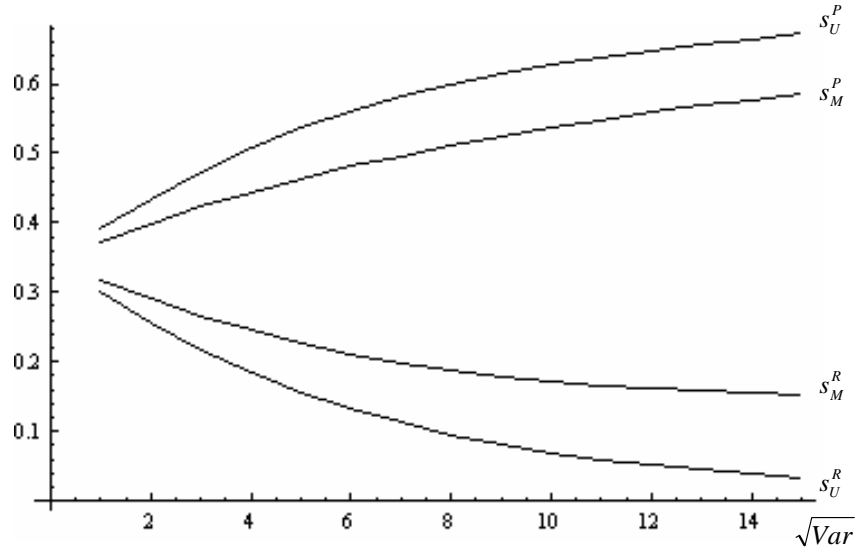


Figure 9: The size of the lobby of the poor s^P and the size of the lobby of the rich s^R as a function of standard deviation \sqrt{Var} for constant mean $\hat{x} = 5.611469$ with $\alpha = 1$: s_U^P , s_U^R (utilitarian government) and s_M^P , s_M^R (pro-median government).

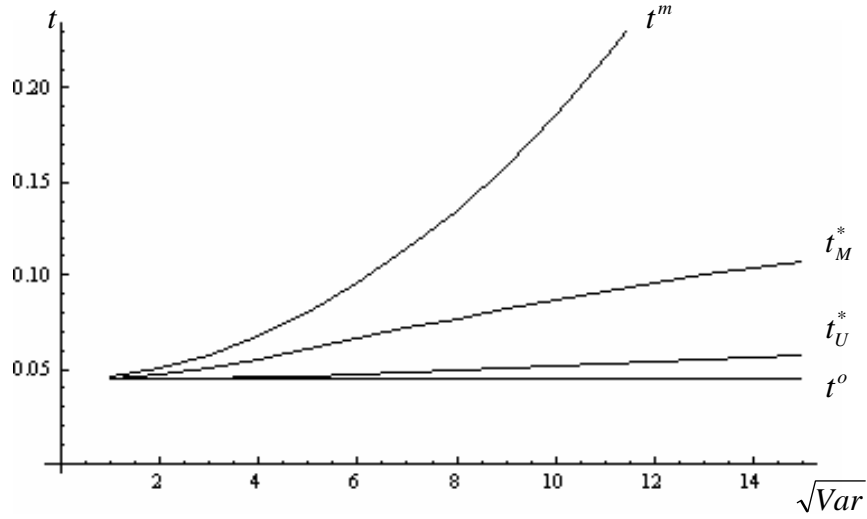


Figure 10: Equilibrium tax rate as a function of standard deviation \sqrt{Var} for constant mean $\hat{x} = 5.611469$ with $\alpha = 1$: t_U^* (utilitarian government) and t_M^* (pro-median government).

CHAPTER III

COMPUTING WELFARE LOSSES FROM DATA UNDER IMPERFECT COMPETITION WITH HETEROGENEOUS GOODS

This chapter is based on Corchón and Zudenkova (2009).

3.1 Introduction

One of the most robust findings of Industrial Organization theory is that market equilibrium very often yields inefficient allocations. But how large are these inefficiencies? This topic has inspired a considerable amount of empirical research, from the paper by Harberger (1954) to the work of Cowling and Mueller (1978), among many others.

In contrast, the theoretical literature is sparse and focuses on the case of homogeneous products. In that case, when demand and costs are linear and firms are identical, it is well known that the percentage of welfare losses (PWL) in a Cournot Equilibrium is $\frac{1}{(1+n)^2}$ where n is the number of firms. McHardy (2000) showed that when demand is quadratic, welfare losses can be 30% larger than in the linear model. Anderson and Renault (2003) calculated PWL for a more general class of demand functions. Johari and Tsitsiklis (2005) showed that if average costs are not increasing and the inverse demand function is concave, PWL is less than $\frac{1}{2n+1}$. Finally, Corchón (2008) offered formulae for PWL under free entry and heterogeneous firms. He showed that PWL can be very large under these conditions. The only paper dealing with heterogeneous products is by Cable et al. (1994), who studied a linear duopoly model.

In this paper we analyze PWL in two models of imperfect competition with heterogeneous products and a representative consumer with quasi-linear preferences: a model with linear demand functions, as per Dixit (1979) and Singh and Vives (1984), and a model with

isoelastic demand functions, as per Spence (1976). In both models, firms produce under constant average costs.

Our first step is to find PWL as a function of the fundamentals, i.e., the parameters of the demand and cost functions. As these parameters cannot be observed, our second step is to obtain PWL as a function of observable variables: price, output, number of firms, etc. Where this is not possible, we introduce items that might be estimated such as elasticity of demand. The goal of our analysis is to study the impact of observable variables on PWL.¹ Even though PWL can be calculated directly from the data on a case-by-case basis, our approach pinpoints the theoretical factors explaining PWL.

We first consider the model with linear demand. Assume that firms and demand functions are identical. We show that, given an observation of price, output, marginal cost and the number of firms, there exist parameters of the demand function that convert this observation into a Cournot or a Bertrand equilibrium such that PWL is arbitrary (Propositions 14 and 15). This result shows that PWL is unrelated to the differences among profit rates, contrary to Harberger's dictum: "The differences among these profit rates, as between industries, give a broad indication of the extent of resource malallocation" (op. cit. p. 79). In our model all firms have the same rate of return on capital but PWL can be high. It seems that Harberger's procedure picks up welfare losses stemming from the failure of markets to equalize profit rates, and not welfare losses from oligopolistic misallocation. The issues are related, but distinct.

Next we show that if the elasticity of demand can be estimated, PWL in a Cournot equilibrium can be computed from observables (Proposition 16). The elasticity of demand does not add any new information in the case of a Bertrand equilibrium because it can be obtained from the markup and the first-order condition of profit maximization. We show that if the cross elasticity of demand can be estimated, PWL can be computed from observations (Proposition 18). Finally, we study how PWL depends on these variables (Propositions 17 and 19). Some results are as expected, but others are not: PWL is

¹This paper does not focus on the maximal PWL. Since product differentiation reduces competition, PWL in our framework is at least as high as it would be under product homogeneity. See Footnote 7 for the case of identical firms and formulae (24), and (26) for the case of non-identical firms.

decreasing on the price-marginal cost margins (often referred to as the "monopoly index", Lerner, 1934), for example, in both Cournot and Bertrand equilibria.² Another surprising result is that PWL increases with the elasticity of demand in a Bertrand equilibrium. Why is this so? Consider two markets, A and B, and let the price-marginal cost margin be larger in A than in B. This means that the triangle that represents welfare losses is larger in A than in B. However, the realized welfare is also larger in A than in B because the demand function in A is above the demand function in B. *A priori*, there is no good reason to expect that one effect is larger than the other. In fact, as we noted before, when costs and demand are linear and firms are identical, these two effects cancel each other out and PWL only depends on the number of firms.³ The same argument goes for demand elasticity: a larger demand elasticity means less welfare losses and less realized welfare, so the total effect is ambiguous.

Next we introduce heterogeneity in demand and costs. We focus on the relationship between concentration and welfare losses. Some papers have found that the Hirschman-Herfindahl (H) index of concentration is not a good measure of welfare losses. Daughety (1990) came to this conclusion because more concentration may be associated with a larger output in a leader-follower equilibrium. In papers by Farrell and Shapiro (1990), Cable et al. (1994) and Corchón (2008), the same result was related to the fact that the firms could be of different sizes.⁴ This finding contrasts with the 1992 Merger Guidelines issued by the Federal Trade Commission (FTC), where H is considered a reasonable measure of welfare losses (Coate, 2005). We show that when it is optimal to allow all firms to produce, PWL increases with H in both Cournot and Bertrand equilibria (Proposition 22). This case arises when goods are poor substitutes. We also show that when it is optimal to allow only one firm to produce, PWL decreases with H. This is what happened in the papers cited above where products are perfect substitutes.

²This was noted by Formby and Leyson (1982) in the case of monopoly.

³In other words, price-marginal cost margins do not control for the size of demand. Thus, a high margin might indicate either that demand is very large and firms are having good times—even if they are very competitive—or that firms are "exploiting" consumers and destroying a large part of the surplus. This is true even if actual production is known, because the price-marginal cost margin is a poor indicator of efficient production.

⁴The point that minor firms may be harmful for welfare was first made by Lahiri and Ono (1988).

Thus, we find that concentration is bad (good) for welfare when goods are poor (good) substitutes. The reason is that efficient production must balance cost savings against consumer satisfaction. The former favors concentrating production in the most efficient firms, while the latter may require considerable diversification of production. If the last effect is not very large (i.e., when the products are close substitutes), cost savings drive efficiency and thus concentration does not harm efficiency. If the products are poor substitutes, however, efficient production requires output dispersion so concentration is harmful. We also show that at the value of H proposed by the FTC as a threshold for a concentrated industry, PWL is large in a Cournot equilibrium but may be small in a Bertrand equilibrium.⁵

In Section 3.3 we assume that the representative consumer has preferences representable by a CES utility function. We also assume a large number of identical firms. This model (Spence, 1976) and its variants (e.g. Dixit and Stiglitz, 1977) are popular in the fields of monopolistic competition, international trade, geography and economics. We depart from these models, however, by assuming that the number of firms is exogenous. The reason for this difference is that to endogenize the number of firms we need fixed costs, which may produce large PWL (Corchón, 2008). Since in this paper we want to focus on the PWL produced by product heterogeneity, we must assume that the number of firms is given. We show that PWL tends to zero as demand elasticity tends to infinity, and that PWL tends to one as the degree of homogeneity of the CES function tends to one (Proposition 25). This result qualifies a conjecture of Stigler (1949): "...the predictions of this standard model of imperfect competition differ only in unimportant respects from those of the theory of competition because the underlying conditions will usually be accompanied by very high demand elasticities for the individual firms". Although a high elasticity of demand makes PWL small in this model, given any elasticity of demand we can obtain a PWL as close to one as we wish.

Next, we show that PWL can be recovered from an observation of the price, output, marginal cost and number of firms (Proposition 26). However, a low price-marginal cost

⁵Despite the fact that, as shown by Amir and Jin (2001), H is always higher in a Bertrand equilibrium than in a Cournot equilibrium.

margin does not guarantee that PWL is small; even if the price tends to the marginal cost, when the number of firms is sufficiently large, PWL may exceed that obtained in a linear model under monopoly. Moreover, when the number of firms tends to infinity, PWL is decreasing in the price-marginal cost margin (Proposition 27). This is another case where price-marginal cost margins and welfare losses are not related in the way we had previously thought.

Summing up, we have three main conclusions. First, our main message is positive: obtaining PWL from data is possible in two well-known models of imperfect competition. Second, the roles of rates of returns, markups and the elasticity of demand on PWL are not always what they have been thought to be. Finally, we explain the role of the H index. Our formulae unify previous views on the role of elasticities, markups and concentration in a precise way, with results that may be useful for policy making.

3.2 *The Linear Model*

In this section we assume that inverse demand is linear and that goods are substitutes.⁶ In the first subsection we assume that all firms are identical, which allows for clean formulae of welfare losses. In the second subsection we study the case where costs and the intercepts of inverse demands vary among firms. The resulting formulae for PWL will then be used to discuss the role of concentration in oligopolistic markets.

3.2.1 *The Symmetric Case*

The market is composed of n firms. The output and price of firm i are denoted by x_i and p_i respectively. The firms are identical, sharing the cost function cx_i . There is a representative consumer with a quadratic utility function

$$U = \alpha \sum_{i=1}^n x_i - \frac{\beta}{2} \sum_{i=1}^n x_i^2 - \frac{\gamma}{2} \sum_{i=1}^n x_i \sum_{j \neq i} x_j + M, \quad \alpha > c, \quad \beta > \gamma \geq 0,$$

where M is the consumption of an outside good which is the numeraire. The budget constraint is $\sum_{i=1}^n p_i x_i + M = I$, where I is a given income. Substituting M into U and eliminating I (which is constant), we obtain $U = \alpha \sum_{i=1}^n x_i - \frac{\beta}{2} \sum_{i=1}^n x_i^2 - \frac{\gamma}{2} \sum_{i=1}^n x_i \sum_{j \neq i} x_j -$

⁶We study the case of complements in a companion working paper (Corchón and Zudenkova, 2008).

$\sum_{i=1}^n p_i x_i$. We call this quantity the consumer surplus. Under our assumptions, this function is concave. The first-order condition (FOC) of consumer surplus maximization yields $p_i = \alpha - \beta x_i - \gamma \sum_{j \neq i} x_j$, $i = 1, 2, \dots, n$. If $\gamma = 0$ the products are independent, while if $\gamma \simeq \beta$ they are almost perfect substitutes.

Definition 8 A linear market is a list $\{\alpha, \beta, \gamma, c, n\}$ with $\alpha > c$, $\beta > \gamma \geq 0$ and $n \in \mathbb{N}$.

Social welfare is defined as

$$W = \alpha \sum_{i=1}^n x_i - \frac{\beta}{2} \sum_{i=1}^n x_i^2 - \frac{\gamma}{2} \sum_{i=1}^n x_i \sum_{j \neq i} x_j - c \sum_{i=1}^n x_i. \quad (11)$$

The social optimum is a list of outputs that maximize social welfare. It is easy to see that the optimal outputs x_i^o (which are all identical) and the social welfare in the optimum W^o are

$$x_i^o = \frac{\alpha - c}{\beta + \gamma(n-1)} \text{ and } W^o = \frac{n(\alpha - c)^2}{2(\beta + (n-1)\gamma)}.$$

Now we are ready to define our equilibrium concepts.

Definition 9 A Cournot equilibrium in a linear market is a list of outputs $(x_1^c, x_2^c, \dots, x_n^c)$ such that for each i , x_i^c maximizes $(\alpha - \beta x_i - \gamma \sum_{j \neq i} x_j^c - c) x_i$.

From the FOC of profit maximization we find that

$$x_i^c = \frac{\alpha - c}{2\beta + \gamma(n-1)}, \quad i = 1, 2, \dots, n. \quad (12)$$

In order to define a Bertrand equilibrium we write the demand for firm i :

$$x_i = \frac{\alpha(\beta - \gamma) - p_i(\beta + \gamma(n-2)) + \gamma \sum_{j \neq i} p_j}{(\beta - \gamma)(\beta + \gamma(n-1))} \equiv x_i^b(p_i, p_{-i}), \quad i = 1, 2, \dots, n, \quad (13)$$

where p_{-i} is a list of all prices minus p_i . Now we can define a Bertrand equilibrium.

Definition 10 A Bertrand equilibrium in a linear market is a list of prices $(p_1^b, p_2^b, \dots, p_n^b)$ such that for each i , p_i^b maximizes $(p_i - c) x_i^b(p_i, p_{-i}^b)$.

From the FOC of profit maximization we obtain

$$p_i^b = \frac{\alpha(\beta - \gamma) + c(\beta + \gamma(n-2))}{2\beta + \gamma(n-3)}, \quad i = 1, 2, \dots, n. \quad (14)$$

Let W^c be social welfare evaluated at the Cournot equilibrium. Let us define the percentage of welfare losses in a Cournot equilibrium as

$$PWL^c \equiv \frac{W^o - W^c}{W^o}.$$

Lemma 11 *In linear markets the percentage of welfare losses in Cournot equilibrium is*

$$PWL^c = \frac{1}{\left(2 + (n-1) \frac{\gamma}{\beta}\right)^2}$$

Proofs of this and other results are given in the Appendix.

Notice that PWL^c is decreasing in the degree of product differentiation, $\frac{\gamma}{\beta}$. Thus, the minimal PWL^c is $\frac{1}{(n+1)^2}$ and occurs when $\gamma \simeq \beta$, i.e., when the products are perfect substitutes. The maximal PWL^c is equal to 0.25 and occurs for the minimal value of $\frac{\gamma}{\beta}$, which is zero when products are independent.

The following Lemma derives PWL^b , the percentage of welfare losses in a Bertrand equilibrium.

Lemma 12 *In linear markets the percentage of welfare losses in Bertrand equilibrium is*

$$PWL^b = \left(\frac{1 - \frac{\gamma}{\beta}}{2 + (n-3) \frac{\gamma}{\beta}} \right)^2. \quad (15)$$

Note that PWL^b is decreasing in the degree of product differentiation $\frac{\gamma}{\beta}$. Thus, minimal PWL^b is (almost) zero and occurs when $\gamma \simeq \beta$, i.e., when the products are perfect substitutes. Maximal PWL^b is equal to 0.25 and occurs for $\frac{\gamma}{\beta} = 0$, when products are independent. Clearly, if $n = 1$, $PWL^j = 0.25$, $j = c, b$, so for the remainder of this section we will assume that $n > 1$.

We are interested in the PWL yielded by imperfectly competitive markets, conditional on certain observable variables: market prices, outputs, marginal cost, and the number of firms. We assume that the marginal cost is observable because under constant returns, the marginal cost equals the average variable cost, which in principle can be observed (wages, raw materials, etc.). Formally:

Definition 13 *An observation is a list $\{\mathbf{p}, \mathbf{x}_i, \mathbf{c}, \mathbf{n}\}$ where \mathbf{p} is the market price, \mathbf{x}_i is the output of firm i , \mathbf{c} ($< \mathbf{p}$) is the marginal cost and \mathbf{n} is the number of firms.*

Let us relate PWL to the observable variables. First consider the Cournot equilibrium.

Proposition 14 *Given an observation $\{\mathbf{p}, \mathbf{r}_i, \mathbf{c}, \mathbf{n}\}$ and a number $\mathbf{v} \in (\frac{1}{(\mathbf{n}+1)^2}, 0.25]$ there is a linear market $\{\alpha, \beta, \gamma, \mathbf{c}, \mathbf{n}\}$ such that $(\mathbf{r}_i, \mathbf{r}_i, \dots, \mathbf{r}_i)$ is a Cournot equilibrium for this market, where $\mathbf{p} = \alpha - \beta \mathbf{r}_i - \gamma(\mathbf{n} - 1)\mathbf{r}_i$ and $PWL^c = \mathbf{v}$.*

Now we turn to the Bertrand equilibrium.

Proposition 15 *Given an observation $\{\mathbf{p}, \mathbf{r}_i, \mathbf{c}, \mathbf{n}\}$ and a number $\mathbf{v} \in (0, 0.25]$ there is a linear market $\{\alpha, \beta, \gamma, \mathbf{c}, \mathbf{n}\}$ such that $(\mathbf{p}, \mathbf{p}, \dots, \mathbf{p})$ is a Bertrand equilibrium for this market, $\mathbf{r}_i = x_i^b(\mathbf{p}, \mathbf{p}_{-i})$, where \mathbf{p}_{-i} is a list of $\mathbf{n} - 1$ identical \mathbf{p} and $PWL^b = \mathbf{v}$.*

Propositions 14 and 15 show that observable variables put very few restrictions on PWL . In particular, neither price-marginal cost margins nor profit rates have any relationship with PWL . Let us look for restrictions that can take a bite out of PWL .⁷ Suppose that the demand elasticity, denoted by ε , is observable. From (13), we have

$$\varepsilon \equiv -\frac{\partial x_i}{\partial p_i} \frac{\mathbf{p}}{\mathbf{r}_i} = \frac{\beta + \gamma(\mathbf{n} - 2)}{(\beta - \gamma)(\beta + \gamma(\mathbf{n} - 1))} \frac{\mathbf{p}}{\mathbf{r}_i}.$$

Let us introduce a new piece of notation: $\mathfrak{T} \equiv \varepsilon \frac{\mathbf{p} - \mathbf{c}}{\mathbf{p}}$. Now we have the following result.

Proposition 16 *Given an observation $\{\mathbf{p}, \mathbf{r}_i, \mathbf{c}, \mathbf{n}, \varepsilon\}$ such that $\mathfrak{T} \equiv \varepsilon \frac{\mathbf{p} - \mathbf{c}}{\mathbf{p}} \geq 1$, there is a linear market $\{\alpha, \beta, \gamma, \mathbf{c}, \mathbf{n}\}$ such that $(\mathbf{r}_i, \mathbf{r}_i, \dots, \mathbf{r}_i)$ is a Cournot equilibrium for this market, $\mathbf{p} = \alpha - \beta \mathbf{r}_i - \gamma(\mathbf{n} - 1)\mathbf{r}_i$ and*

$$PWL^c = \frac{1}{\left(2 + \frac{(\mathfrak{T} - 1)(\mathbf{n} - 2) + \sqrt{(\mathfrak{T} - 1)(\mathbf{n}^2 \mathfrak{T} - (\mathbf{n} - 2)^2)}}{2\mathfrak{T}}\right)^2} \quad (16)$$

According to Proposition 16 we can calculate PWL^c from three variables: the number of firms, the elasticity of demand and the price-marginal cost ratio. Let us study how PWL^c depends on \mathbf{n} and \mathfrak{T} . Notice that observable variables are not independent and that in general, a variation in just one observable variable cannot be obtained by a variation

⁷The maximum PWL in both Cournot and Bertrand equilibria occurs when $\gamma \simeq 0$, namely $PWL \simeq 0.25$, which corresponds to PWL under monopoly.

of a single unobservable variable (see the first three equations of the proof of Proposition 16). Our exercise just gives us the difference in PWL between two markets in which all observables except one are identical. Thus, it emphasizes the role played by the observables, which sometimes contradicts intuition.

Proposition 17 *PWL^c is decreasing in \mathbf{n} , the elasticity of demand, and the price-marginal costs margins.*

In Proposition 17, the signs of the effects of the number of firms and demand elasticity are just as expected: more competition—i.e., a higher value of \mathbf{n} or ε —is good. However the effect of price-marginal cost margins runs counter to intuition. As we remarked in the introduction, this is because the price-marginal cost margin affects both welfare losses and realized welfare.

Note that we have been applying comparative statics, treating the observable variables \mathbf{n} , ε and $\frac{\mathbf{p}-\mathbf{c}}{\mathbf{p}}$ as exogenous. This approach provides policy-makers with a tool to predict changes in PWL due to changes in just one of the observables, taking all other variables as given. If both demand elasticity and the price-marginal cost margin change, one needs to consider the comparative statics of PWL with respect to the factor $\mathfrak{T} \equiv \varepsilon \frac{\mathbf{p}-\mathbf{c}}{\mathbf{p}}$ given in (31).

We now consider the Bertrand equilibrium. In this case, the FOC of profit maximization can be written as $p_i = \varepsilon(p_i - c)$. Thus, an observation of ε does not add any new information once p_i and c are observed. A way out of this problem is provided if the cross elasticity of demand $\frac{\partial x_i}{\partial p_j} \frac{p_j}{x_i}$, denoted by ρ , is observable.

Proposition 18 *Given an observation $\{\mathbf{p}, \mathbf{r}_i, \mathbf{c}, \mathbf{n}, \rho\}$ such that $\frac{\mathbf{p}}{\mathbf{p}-\mathbf{c}} > \rho(\mathbf{n}-1) \geq 0$, there is a linear market $\{\alpha, \beta, \gamma, \mathbf{c}, \mathbf{n}\}$ such that $(\mathbf{p}, \mathbf{p}, \dots, \mathbf{p})$ is a Bertrand equilibrium for this market, $\mathbf{r}_i = x_i^b(\mathbf{p}, \mathbf{p}_{-i})$, where \mathbf{p}_{-i} is a list of $\mathbf{n}-1$ identical \mathbf{p} and*

$$PWL^b = \left(\frac{\frac{\mathbf{p}}{\mathbf{p}-\mathbf{c}} - \rho(\mathbf{n}-1)}{2\frac{\mathbf{p}}{\mathbf{p}-\mathbf{c}} - \rho(\mathbf{n}-1)} \right)^2. \quad (17)$$

The formula (17) allows for the calculation of PWL in a Bertrand equilibrium from just three magnitudes: the number of firms, the price-marginal cost margins (or alternatively,

the elasticity of demand), and the cross elasticity of demand. Let us analyze the impact of a change in observable variables on PWL^b .

Proposition 19 *PWL^b is decreasing in the number of firms, the price-marginal cost margins, and the cross elasticity of demand. PWL^b is increasing in the elasticity of demand.*

Proposition 19 confirms our intuitions about the role of the number of firms and the cross elasticity of demand on welfare losses, namely that an increase in the number of firms decreases PWL and an increase in the cross elasticity of demand decreases PWL. The impacts of the price-marginal cost margin and demand elasticity, on the other hand, are contrary to intuition. Again, we have to bear in mind that these two variables affect both welfare losses and realized welfare.

3.2.2 Heterogeneous Firms

We now extend the model presented in Subsection 3.2.1 to the case where firms are heterogeneous on two counts. The marginal costs c_i may be different for each firm i . The parameter α , now denoted α_i , may also vary across firms.⁸ Assume $\alpha_i > c_i$ for all i . The consumer surplus is now

$$U = \sum_{i=1}^n \alpha_i x_i - \frac{\beta}{2} \sum_{i=1}^n x_i^2 - \frac{\gamma}{2} \sum_{i=1}^n x_i \sum_{j \neq i} x_j - \sum_{i=1}^n p_i x_i, \quad \beta > \gamma \geq 0$$

The restrictions below guarantee that the outputs of all firms are positive in Cournot and Bertrand equilibria.

$$2\beta + \gamma(n-1) > \frac{\gamma \sum_{i=1}^n (\alpha_i - c_i)}{\alpha_i - c_i}, \quad i = 1, 2, \dots, n. \quad (18)$$

$$\frac{(\beta + \gamma(n-1))(2\beta + \gamma(n-3))}{\beta + \gamma(n-2)} > \frac{\gamma \sum_{i=1}^n (\alpha_i - c_i)}{\alpha_i - c_i} \quad i = 1, 2, \dots, n. \quad (19)$$

Under our assumptions, $U(\cdot)$ is concave. The FOC of consumer surplus maximization yields $p_i = \alpha_i - \beta x_i - \gamma \sum_{j \neq i} x_j$, $i = 1, 2, \dots, n$. Social welfare is now

$$W = \sum_{i=1}^n \alpha_i x_i - \frac{\beta}{2} \sum_{i=1}^n x_i^2 - \frac{\gamma}{2} \sum_{i=1}^n x_i \sum_{j \neq i} x_j - \sum_{i=1}^n c_i x_i. \quad (20)$$

⁸This model has been used, among others, by Häckner (2000) and Hsu and Wang (2005).

Evaluating social welfare in the optimum is not straightforward, because it depends on the number of active firms in the optimum. For the time being, let us assume that the optimal number of active firms is m . Then the optimal outputs, denoted by x_i^o , are equal to

$$x_i^o = \frac{\alpha_i - c_i}{\beta - \gamma} - \frac{\gamma \sum_{i=1}^m (\alpha_i - c_i)}{(\beta + \gamma(m-1))(\beta - \gamma)}, \quad i = 1, 2, \dots, m \quad (21)$$

and the aggregate output in the optimum, denoted by x^o , is equal to

$$x^o = \sum_{i=1}^m x_i^o = \frac{\sum_{i=1}^m (\alpha_i - c_i)}{\beta + \gamma(m-1)}$$

We now find the optimal number of firms m . Let us rank firms according to the value of $\alpha_i - c_i$. Without loss of generality assume that $\alpha_v - c_v \geq \alpha_{v+1} - c_{v+1}$, $v = 1, 2, \dots, n-1$. Clearly, if firm v produces a positive output in the optimum, firms $v-1, v-2$, etc. also produce positive outputs in the optimum. Suppose that it is optimal for firms 1 through $k-1$ to produce positive outputs. By evaluating $\frac{\partial W}{\partial x_k}$ in (20) at $x_k = 0$ and $x_j = x_j^o$, $j = 1, \dots, k-1$ according to (21), we obtain

$$\frac{\partial W}{\partial x_k} = \alpha_k - c_k - \gamma \sum_{j=1}^{k-1} x_j^o. \quad (22)$$

If $\frac{\partial W}{\partial x_k} \leq 0$, clearly, $x_k^o = 0$. If $\frac{\partial W}{\partial x_k} > 0$, firm k must produce a positive output in the optimum.

This algorithm requires knowledge of all the parameters defining a market. In a companion working paper (Corchón and Zudenkova, 2008), we show that all these parameters can be recovered from market data and demand elasticities by a method identical to that applied in Propositions 16 and 18. We will focus on two particular cases. First, when $\beta(\alpha_2 - c_2) \leq \gamma(\alpha_1 - c_1)$, only firm 1 will produce a positive output in the optimum since from (21) and (22), $\frac{\partial W}{\partial x_2} = \alpha_2 - c_2 - \gamma \frac{\alpha_1 - c_1}{\beta} \leq 0$. Second, when

$$(\alpha_n - c_n)(\beta + \gamma(n-2)) > \gamma \sum_{i=1}^{n-1} (\alpha_i - c_i), \quad (23)$$

the number of active firms is the same in optimum and in equilibrium, because from (21) and (22), $\frac{\partial W}{\partial x_n} = \alpha_n - c_n - \gamma \sum_{j=1}^{n-1} x_j^o > 0$. Notice that the conditions in (18) and (19) are implied by (23).

In this framework, a Cournot equilibrium is a list of outputs $(x_1^c, x_2^c, \dots, x_n^c)$ such that for each i , x_i^c maximizes $(\alpha_i - \beta x_i - \gamma \sum_{j \neq i} x_j^c - c_i) x_i$. From the FOC of profit maximization, we obtain

$$x_i^c = \frac{\alpha_i - c_i}{2\beta - \gamma} - \frac{\gamma}{2\beta - \gamma} \frac{\sum_{i=1}^n (\alpha_i - c_i)}{2\beta + \gamma(n-1)}.$$

And the aggregate output at the Cournot equilibrium is

$$x^c = \sum_{i=1}^n x_i^c = \frac{\sum_{i=1}^n (\alpha_i - c_i)}{2\beta + \gamma(n-1)}.$$

In order to compute a Bertrand equilibrium we first write the demand for firm i :

$$x_i = \frac{\alpha_i (\beta + \gamma(n-2)) - p_i (\beta + \gamma(n-2)) - \gamma \sum_{j \neq i} (\alpha_j - p_j)}{(\beta - \gamma)(\beta + \gamma(n-1))} \equiv x_i^b(p_i, p_{-i}).$$

A Bertrand equilibrium is a list $(p_1^b, p_2^b, \dots, p_n^b)$ such that for all i p_i^b maximizes $(p_i - c_i) x_i^b(p_i, p_{-i}^b)$.

Then we have

$$x_i^b = \frac{\beta + \gamma(n-2)}{(\beta - \gamma)(2\beta + \gamma(2n-3))} \left(\alpha_i - c_i - \gamma \frac{\beta + \gamma(n-2)}{(2\beta + \gamma(n-3))(\beta + \gamma(n-1))} \sum_{i=1}^n (\alpha_i - c_i) \right),$$

and the aggregate output at the Bertrand equilibrium is

$$x^b = \frac{\beta + \gamma(n-2)}{(2\beta + \gamma(n-3))(\beta + \gamma(n-1))} \sum_{i=1}^n (\alpha_i - c_i).$$

Next, we link PWL to the Hirschman-Herfindahl index of concentration. Let s_i^j be the market share of firm i in a Cournot equilibrium ($j = c$), a Bertrand equilibrium ($j = b$), or in the optimum ($j = o$). We define the Hirschman-Herfindahl index of concentration in either equilibrium as $H^j \equiv \sum_{i=1}^n (s_i^j)^2$, $j = c, b$. In the optimum, we define it as $H^o \equiv \sum_{i=1}^m (s_i^o)^2$.⁹ Amir and Jin (2001) show that in our framework $H^b > H^c$.

Lemma 20 *With heterogeneous firms the percentage of welfare losses in Cournot equilibrium is*

$$PWL^c = 1 - \left(\frac{1 + \frac{\gamma}{\beta}(m-1)}{m \frac{\gamma}{\beta} + \left(2 - \frac{\gamma}{\beta}\right) \sum_{i=1}^m s_i^c} \right)^2 \frac{H^c \left(3 - \frac{\gamma}{\beta}\right) + \frac{\gamma}{\beta}}{H^o \left(1 - \frac{\gamma}{\beta}\right) + \frac{\gamma}{\beta}}. \quad (24)$$

⁹Notice that in our framework products can be imperfect substitutes, so interpreting $\sum_{i=1}^n (s_i)^2$ as the Hirschman-Herfindahl index of concentration may be a bit problematic.

Note that PWL^c here depends on the degree of product differentiation $\frac{\gamma}{\beta}$, the number of active firms in the optimum m , the sum of the market shares of the m largest firms $\sum_{i=1}^m s_i^c$, and the Hirschman-Herfindahl indices of concentration H^c and H^o evaluated at the Cournot equilibrium and optimum respectively. When $m = 1$ we have

$$PWL^c(m = 1) = 1 - \frac{H^c \left(3 - \frac{\gamma}{\beta}\right) + \frac{\gamma}{\beta}}{\left(\frac{\gamma}{\beta} + \left(2 - \frac{\gamma}{\beta}\right) s_1^c\right)^2},$$

which is decreasing in H^c . In the other extreme case where $m = n$ —i.e., the number of active firms is the same in the optimum and at a Cournot equilibrium—we prove that

$$PWL^c(m = n) = \frac{H^c \left(1 + (n-1) \frac{\gamma}{\beta}\right) - \frac{\gamma}{\beta}}{H^c \left(2 - \frac{\gamma}{\beta}\right)^2 \left(1 + (n-1) \frac{\gamma}{\beta}\right) + \left(\frac{\gamma}{\beta}\right)^2 \left(n - 2 - (n-1) \frac{\gamma}{\beta}\right)}. \quad (25)$$

If all firms are identical, $H^c = \frac{1}{n}$ and $PWL^c(m = n) = \frac{1}{\left(2 + (n-1) \frac{\gamma}{\beta}\right)^2}$, as in Lemma 11. Notice that H^c and $\frac{\gamma}{\beta}$ are less than one, so for reasonable values of n it makes sense to evaluate (25) as if n were a large number. In this case (25) simplifies to

$$PWL^c(m = n, n \text{ large}) = \frac{H^c}{H^c \left(2 - \frac{\gamma}{\beta}\right)^2 + \frac{\gamma}{\beta} \left(1 - \frac{\gamma}{\beta}\right)}, \text{ and}$$

$$\frac{\partial PWL^c(m = n, n \text{ large})}{\partial \frac{\gamma}{\beta}} = - \frac{H^c \left(1 - 2\frac{\gamma}{\beta} - 2H^c(2 - \frac{\gamma}{\beta})\right)}{\left(H^c \left(2 - \frac{\gamma}{\beta}\right)^2 + \frac{\gamma}{\beta} \left(1 - \frac{\gamma}{\beta}\right)\right)^2}.$$

The latter is negative for $\frac{\gamma}{\beta} \in (0, \frac{1-4H^c}{2(1-H^c)})$ and positive for $\frac{\gamma}{\beta} \in (\frac{1-4H^c}{2(1-H^c)}, 1)$. Thus, the minimum occurs at $\frac{\gamma}{\beta} = \frac{1-4H^c}{2(1-H^c)}$. When $H^c = 0.18$, which the FTC considers the threshold for a concentrated industry, the minimal PWL^c is 0.241967. This value is a large lower bound.

Now we consider welfare losses in a Bertrand equilibrium.

Lemma 21 *In a Bertrand equilibrium with heterogeneous firms*

$$PWL^b = 1 - \left(\frac{\left(1 + \frac{\gamma}{\beta}(n-2)\right) \left(1 + \frac{\gamma}{\beta}(m-1)\right)}{m \frac{\gamma}{\beta} \left(1 + \frac{\gamma}{\beta}(n-2)\right) + \left(1 - \frac{\gamma}{\beta}\right) \left(2 + \frac{\gamma}{\beta}(2n-3)\right) \sum_{i=1}^m s_i^b} \right)^2. \quad (26)$$

$$\frac{H^b \left(1 - \frac{\gamma}{\beta}\right) \left(3 + \frac{\gamma}{\beta}(3n-4)\right) + \frac{\gamma}{\beta} \left(1 + \frac{\gamma}{\beta}(n-2)\right)}{\left(H^o \left(1 - \frac{\gamma}{\beta}\right) + \frac{\gamma}{\beta}\right) \left(1 + \frac{\gamma}{\beta}(n-2)\right)}$$

Thus, PWL^b depends on the degree of product differentiation $\frac{\gamma}{\beta}$, the number of active firms in the optimum m and in a Bertrand equilibrium n , the sum of the market shares of the m largest firms $\sum_{i=1}^m s_i^b$, and the Hirschman-Herfindahl indices of concentration H^b and H^o , evaluated at a Bertrand equilibrium and in the optimum respectively.

As before, let us consider two special cases. First, when in the optimum only firm 1 is used by the planner. Thus, $m = 1$ and

$$PWL^b(m = 1) = 1 - \left(\frac{1 + \frac{\gamma}{\beta}(n-2)}{\frac{\gamma}{\beta}(1 + \frac{\gamma}{\beta}(n-2)) + (1 - \frac{\gamma}{\beta})(2 + \frac{\gamma}{\beta}(2n-3))s_1^b} \right)^2 \frac{H^b(1 - \frac{\gamma}{\beta})(3 + \frac{\gamma}{\beta}(3n-4)) + \frac{\gamma}{\beta}(1 + \frac{\gamma}{\beta}(n-2))}{1 + \frac{\gamma}{\beta}(n-2)}.$$

For $\beta \simeq \gamma$, this formula reduces to $PWL^b(m = 1) = 0$, as one expects for a Bertrand equilibrium under product homogeneity. Notice that $PWL^b(m = 1)$ is decreasing in H^b . Second, when the number of active firms is the same in the optimum and in a Bertrand equilibrium, after lengthy calculations, we obtain the result

$$PWL^b(m = n) = \frac{(H^b(1 + \frac{\gamma}{\beta}(n-1)) - \frac{\gamma}{\beta})(1 - \frac{\gamma}{\beta})(1 + \frac{\gamma}{\beta}(n-1))}{H^b(1 - \frac{\gamma}{\beta})(2 + \frac{\gamma}{\beta}(2n-3))^2 + (\frac{\gamma}{\beta})^2(n-2 + \frac{\gamma}{\beta}(3 + (n-3)n))}. \quad (27)$$

If all firms are identical, $H^b = \frac{1}{n}$ and $PWL^b(m = n) = \left(\frac{1 - \frac{\gamma}{\beta}}{2 + \frac{\gamma}{\beta}(n-3)} \right)^2$ as in Lemma 12. Finally, when n is large, (27) simplifies to

$$PWL^b(m = n, n \text{ large}) = \frac{H^b(1 - \frac{\gamma}{\beta})}{\frac{\gamma}{\beta} + 4H^b(1 - \frac{\gamma}{\beta})},$$

which is decreasing in the degree of product differentiation $\frac{\gamma}{\beta}$. Its maximal value is 0.25 (for $\frac{\gamma}{\beta} = 0$). For $H^b = 0.18$, $PWL^b(m = n, n \text{ large}) = \frac{0.18 - 0.18\frac{\gamma}{\beta}}{0.28\frac{\gamma}{\beta} + 0.72}$ which for values of $\frac{\gamma}{\beta}$ larger than 0.75 is less than 4.8%. So in this case a high concentration does not imply large welfare losses.

From (25) and (27) we obtain the following result:

Proposition 22 $PWL^j(m = n)$ is increasing in H^j , $j = c, b$.

Thus, for $m = n$ PWL increases with H , contrary to what happens when $m = 1$ in both Cournot and Bertrand equilibria. This is because the condition $m = n$ ($m = 1$) is related to goods being poor (good) substitutes. Finally, a word of caution: in Proposition 22 we have assumed that H is independent of all other variables affecting PWL, for example $\frac{\gamma}{\beta}$ and n .

But H does depend on these variables. Strictly speaking, Proposition 22 only applies to variations in H that are caused by variations in the α 's and c 's.

3.3 A Model of a Large Group

In this section we consider the market for a differentiated good supplied by a large number of firms. Typical examples include restaurants, wine, beer, etc. We will not consider entry and fixed costs, which as shown in Corchón (2008) might produce a very high PWL and bias our estimates. As the purpose of this paper is to study the impact of product differentiation alone, we discard these costs. As we will see, even so this model is capable of producing a very high PWL. The model can be interpreted as a monopolistic competition model in which the long-run aspects are not considered. In this framework the relative size of firms is not an important issue, so we will assume that all firms are identical. Also, for convenience, we will assume that firms compete in quantities.

The consumer surplus is given by

$$U = \left(\sum_{i=1}^n x_i^\delta \right)^{\frac{r}{\delta}} - \sum_{i=1}^n p_i x_i, \quad \delta, r \in (0, 1),$$

(see Spence, 1976). The inverse demand function of firm i is $p_i = r \left(\sum_{i=1}^n x_i^\delta \right)^{\frac{r}{\delta}-1} x_i^{\delta-1}$.

Definition 23 A CES Market is a list $\{\delta, r, c, n\}$ with $\delta, r \in (0, 1)$, $c > 0$, and $n \in \mathbb{N}$.

The profit function for firm i is $\pi_i = r \left(\sum_{i=1}^n x_i^\delta \right)^{\frac{r}{\delta}-1} x_i^\delta - c x_i$. Because there are many firms, each firm takes $\sum_{i=1}^n x_i^\delta$ as given. The elasticity of demand, denoted by ϵ , is defined as the inverse of the elasticity of inverse demand:

$$\epsilon = \frac{1}{1 - \delta}. \quad (28)$$

Thus, as $\delta \rightarrow 1$ the elasticity of demand becomes infinite. Now we have the following preliminary result.

Lemma 24 In a CES market

$$PWL^s = 1 - \delta^{\frac{1}{1-r}} \frac{\frac{1}{\delta} - r}{1 - r}. \quad (29)$$

At first glance it is surprising that PWL^s does not depend on the number of firms n . However, we did assume that the number of firms is large. Thus, (29) can be understood as the limit formula when n is large. The following properties of PWL^s are easily proven:

Proposition 25 *i) PWL^s is decreasing in δ .*

ii) $\lim_{\delta \rightarrow 1} PWL^s = 0$ and $\lim_{\delta \rightarrow 0} PWL^s = 1$.

iii) PWL^s is increasing in r .

iv) $\lim_{r \rightarrow 1} PWL^s = 1$ and $\lim_{r \rightarrow 0} PWL^s = 0$.

The explanation of ii) is that when δ is close to one (resp. zero), the products are close to being homogeneous (resp. very differentiated), and welfare losses are small (resp. large); see formula (28). The explanation of iii) is that as r increases (resp. decreases), the gap between the optimal and the equilibrium output increases (resp. decreases); see formulae (33) and (34). It follows from ii) and iv) that it is possible to have a market where the elasticity of demand is close to infinity (i.e., δ is close to 1) and PWL is as close to 1 as we wish.¹⁰ In brief, elasticity of demand is only a partial measure of PWL in this model.

Let us relate PWL^s to the observable variables listed in Definition 13 of the previous section. The FOC of profit maximization imply that $\epsilon = \frac{p}{p-c}$, so in this framework, as in the Bertrand case in the previous section, knowledge of the elasticity of demand is of no help. We will assume that $c(\ln n + \ln p) < p \ln n$; this condition ensures that $r < 1$.

In our construction, the function $ProductLog(t)$ will play a prominent role. This function, called Lambert's W-function, gives the solution for w in $t = we^w$ and has the following properties:¹¹

- i) $ProductLog(t) \in \mathbb{R}$ for $t \in [-\frac{1}{e}, \infty)$;
- ii) $ProductLog(-\frac{1}{e}) = -1$;
- iii) $\lim_{t \rightarrow \infty} ProductLog(t) = \infty$;
- iv) $ProductLog(0) = 0$;
- v) $ProductLog(t)$ is increasing in $t \in [-\frac{1}{e}, \infty)$;
- vi) $e^{a ProductLog(t)} (ProductLog(t))^a = t^a$.

¹⁰ Even if $\delta = r$, $\lim_{\delta \rightarrow 1} PWL^s = 0.2642$, a large number.

¹¹ See Weisstein (1999).

Now we arrive at the main result of this section:

Proposition 26 *Given an observation $\{\mathfrak{p}, \mathfrak{r}_i, \mathfrak{c}, \mathfrak{n}\}$ there is a CES market $\{\delta, r, \mathfrak{c}, \mathfrak{n}\}$ such that $(\mathfrak{p}, \mathfrak{r}_i)$ is an equilibrium for this market, and*

$$PWL^s = 1 - \left(\frac{\mathfrak{c}}{\mathfrak{p}}\right)^{\frac{1}{1-r}} \frac{\frac{\mathfrak{p}}{\mathfrak{c}} - r}{1-r} \text{ where } r = \frac{\text{ProductLog}\left(\mathfrak{n}\mathfrak{p}\mathfrak{r}_i\left(\frac{\mathfrak{p}}{\mathfrak{c}}\ln\mathfrak{n} + \ln\mathfrak{r}_i\right)\right)}{\frac{\mathfrak{p}}{\mathfrak{c}}\ln\mathfrak{n} + \ln\mathfrak{r}_i} \quad (30)$$

An important consequence of Proposition 26 is that, given an observation, there is a unique value of PWL^s . In this case, the number of parameters to be "recovered" is equal to the number of data.

Next, we analyze the properties of PWL^s in (30):

Proposition 27 *The percentage of welfare losses in the CES model is such that*

- i) $\lim_{\mathfrak{n} \rightarrow \infty} PWL^s = 1 - \left(\frac{\mathfrak{c}}{\mathfrak{p}}\right)^{\frac{1}{1-\frac{\mathfrak{c}}{\mathfrak{p}}}} \left(\frac{\mathfrak{p}}{\mathfrak{c}} + 1\right).$
- ii) $\lim_{\frac{\mathfrak{c}}{\mathfrak{p}} \rightarrow 1} PWL^s = 0.$
- iii) $\lim_{\frac{\mathfrak{c}}{\mathfrak{p}} \rightarrow 1} (\lim_{\mathfrak{n} \rightarrow \infty} PWL^s) = 1 - \frac{2}{e} \simeq 0.2642.$

Note that when a finite number of firms are pricing at the marginal cost, PWL^s is close to zero. When an infinite number of firms are pricing at the marginal cost, however, PWL^s is quite high. In fact, it could be argued that formula i) above should be used since we assumed that \mathfrak{n} was large. In this case, PWL^s is decreasing in the price-marginal cost margin, $\frac{\mathfrak{p}-\mathfrak{c}}{\mathfrak{p}}$.¹²

3.4 Conclusion

In this paper we have studied the relationship between observable variables and welfare losses, taking the behavior of firms as given.¹³ Our main message is positive in that relating welfare losses to observables is a feasible endeavor in the models considered in this paper.¹⁴ But there is an important caveat: the calculus of welfare losses depends on the forms of

¹²When n is not large, we have an example, available upon request, showing that PWL is not monotonic in the price-marginal cost margin.

¹³See Sutton (1998) for an approach where the only source of variation between firms is the degree of competitiveness.

¹⁴The models presented in this paper have been selected for their impact in the profession. Thus, the papers by Dixit, Singh and Vives and Spence obtained in the aggregate nearly 1800 citations in Google Scholar.

demand and costs not only at the equilibrium point, but at all points in the domain of these functions. This contrasts with the typical linearization around the equilibrium where a linear form is supposed to represent the characteristics of a general function around the equilibrium point. It is clear that welfare losses depend in a fundamental way on the functional forms, so this issue cannot be dodged. Thus, one can interpret our results by saying that they show welfare consequences of assumptions that are often made about the functional forms of demand and costs.

We end this paper by giving some hints as to how data and elasticities may help us to discriminate among the models. The clearest case is a Bertrand equilibrium. A necessary condition for this equilibrium to be supported by the data is that for all i , $p_i = \varepsilon(p_i - c_i)$ (irrespective of whether the market is linear). If the elasticity of demand cannot be estimated, Proposition 15 says that any observation can be interpreted as a Bertrand equilibrium. The case for the CES model relies on two assumptions. On the one hand, the elasticity of demand must be constant. On the other hand, the cross elasticity of demand (calculated as $\frac{1}{\frac{\partial p_i}{\partial x_j} \frac{x_j}{p_i}}$) should be very high (it amounts to $\frac{n}{r-\delta}$). Finally, let us consider the Cournot equilibrium. Let $\xi \equiv -\frac{\partial p_i}{\partial x_i} \frac{x_i}{p_i}$ be the elasticity of the inverse demand function. From the FOC of profit maximization, we have $\xi p_i = p_i - c_i$. The elasticity ξ can be obtained by inverting the system of demand functions. For instance, in the symmetric case with $n = 2$, $\xi = \frac{\varepsilon}{\varepsilon^2 - \rho^2}$, which when plugged into the FOC yields $\frac{p_i - c_i}{p_i} = \frac{\varepsilon}{\varepsilon^2 - \rho^2}$. If markups, demand and cross elasticities can be estimated, they must obey the previous equation in a Cournot equilibrium.

3.5 Appendix

Proof of Lemma 11: From (11), social welfare in a Cournot equilibrium can be written as $W^c = \alpha n x_i^c - \frac{\beta}{2} n x_i^{c2} - \frac{\gamma}{2} n(n-1) x_i^{c2} - c n x_i^c$. Thus, from (12) we obtain

$$W^c = \frac{n(\alpha - c)^2(3\beta + (n-1)\gamma)}{2(2\beta + (n-1)\gamma)^2}.$$

Then,

$$PWL^c = 1 - \frac{W^c}{W^o} = \frac{1}{\left(2 + (n-1)\frac{\gamma}{\beta}\right)^2}. \blacksquare$$

Proof of Lemma 12: From (14) we find that all firms produce the same output x_i^b , namely

$$x_i^b = \frac{(\alpha - c)(\beta + (n-2)\gamma)}{(2\beta + (n-3)\gamma)(\beta + (n-1)\gamma)}.$$

Social welfare in a Bertrand equilibrium is $W^b = \alpha n x_i^b - \frac{\beta}{2} n x_i^{b2} - \frac{\gamma}{2} n(n-1) x_i^{b2} - c n x_i^b$, or

$$W^b = \frac{n(\alpha - c)^2(3\beta + (n-4)\gamma)(\beta + (n-2)\gamma)}{2(2\beta + (n-3)\gamma)^2(\beta + (n-1)\gamma)}.$$

Thus,

$$PWL^b = 1 - \frac{W^b}{W^o} = \left(\frac{1 - \frac{\gamma}{\beta}}{2 + (n-3)\frac{\gamma}{\beta}} \right)^2. \blacksquare$$

Proof of Proposition 14: Let

$$\alpha = c + \frac{p - c}{\sqrt{v}}, \quad \beta = \frac{p - c}{r_i} \quad \text{and} \quad \gamma = \frac{(p - c)(1 - 2\sqrt{v})}{(n-1)r_i\sqrt{v}}.$$

Clearly, $\alpha > c$ and $\beta > \gamma \geq 0$ since $p > c$, $v > \frac{1}{(n+1)^2}$ and $v \leq 0.25$. We easily see that the linear market $\{\alpha, \beta, \gamma, c, n\}$ yields an equilibrium where $x_i^c = r_i$, $i = 1, 2, \dots, n$, $p = \alpha - \beta r_i - \gamma(n-1)r_i$ and $PWL^c = v$, so the proof is complete. \blacksquare

Proof of Proposition 15: Let

$$\alpha = c + \frac{p - c}{\sqrt{v}}, \quad \beta = \frac{p - c}{r_i} \frac{(\sqrt{v} - 1)(1 + \sqrt{v}(n-3))}{v + n(v - \sqrt{v})} \quad \text{and} \quad \gamma = \frac{c - p}{r_i} \frac{1 - 3\sqrt{v} + 2v}{v + n(v - \sqrt{v})}.$$

It is easy to check that $0 < c < \alpha$ and $\beta > \gamma \geq 0$. The linear market $\{\alpha, \beta, \gamma, c, n\}$ yields a Bertrand equilibrium with $p_i^b = p$, $x_i^b = r_i$ and $PWL^b = v$, so the proof is complete. \blacksquare

Proof of Proposition 16: Let

$$\begin{aligned} \alpha &= c + \frac{p}{2\varepsilon} \left(\mathfrak{T}(n+2) - (n-2) + \sqrt{(\mathfrak{T}-1)(n^2\mathfrak{T} - (n-2)^2)} \right) \\ \beta &= \frac{p - c}{r_i} \\ \gamma &= \frac{(\mathfrak{T}-1)(n-2) + \sqrt{(\mathfrak{T}-1)(n^2\mathfrak{T} - (n-2)^2)}}{2\mathfrak{T}(n-1)} \frac{p - c}{r_i}. \end{aligned}$$

Clearly, $\beta > 0$. We need to show that $0 \leq \frac{\gamma}{\beta} < 1$ and $\alpha > c$. For $\mathfrak{T} \geq 1$ the square root is defined in real numbers and $\sqrt{(\mathfrak{T}-1)(n^2\mathfrak{T} - (n-2)^2)} \geq (\mathfrak{T}-1)(n-2)$ because if not, $n^2\mathfrak{T} < (n-2)^2\mathfrak{T}$, which is impossible. Then the condition $0 \leq \frac{\gamma}{\beta} < 1$ amounts to

$$0 \leq \frac{(\mathfrak{T}-1)(n-2) + \sqrt{(\mathfrak{T}-1)(n^2\mathfrak{T} - (n-2)^2)}}{2\mathfrak{T}(n-1)} < 1 \implies 4\mathfrak{T}(n-1)^2 > 0,$$

which always holds for $\mathfrak{T} \in [1, \infty)$. The condition $\alpha > \mathfrak{c}$ amounts to $(\mathfrak{T} - 1)(\mathfrak{n} - 2) + \sqrt{(\mathfrak{T} - 1)(\mathfrak{n}^2\mathfrak{T} - (\mathfrak{n} - 2)^2)} + 4\mathfrak{T} > 0$, which holds for $\mathfrak{T} \in [1, \infty)$. It is straightforward to show that the linear market $\{\alpha, \beta, \gamma, \mathfrak{c}, \mathfrak{n}\}$ yields a Cournot equilibrium where $x_i^c = \mathfrak{x}_i$, $\mathfrak{p} = \alpha - \beta\mathfrak{x}_i - \gamma(\mathfrak{n} - 1)\mathfrak{x}_i$, and PWL^c as defined by (16). ■

Proof of Proposition 17: From (16) we find

$$\frac{\partial PWL^c}{\partial \mathfrak{n}} = -\frac{8\mathfrak{T}^2 \left(\mathfrak{T} - 1 + \frac{(2 + \mathfrak{n}(\mathfrak{T} - 1))(\mathfrak{T} - 1)}{\sqrt{(\mathfrak{T} - 1)(\mathfrak{n}^2\mathfrak{T} - (\mathfrak{n} - 2)^2)}} \right)}{\left(2 + \mathfrak{n}(\mathfrak{T} - 1) + 2\mathfrak{T} + \sqrt{(\mathfrak{T} - 1)(\mathfrak{n}^2\mathfrak{T} - (\mathfrak{n} - 2)^2)} \right)^3} < 0.$$

Next we compute $\frac{\partial \left(\frac{\gamma}{\beta} \right)}{\partial \mathfrak{T}}$:

$$\frac{\partial \left(\frac{\gamma}{\beta} \right)}{\partial \mathfrak{T}} = \frac{\mathfrak{n}(4 + \mathfrak{n}(\mathfrak{T} - 1) - 2\mathfrak{T}) - 2(2 - \mathfrak{T}) + (\mathfrak{n} - 2)\sqrt{(\mathfrak{T} - 1)(\mathfrak{n}^2\mathfrak{T} - (\mathfrak{n} - 2)^2)}}{2(\mathfrak{n} - 1)\mathfrak{T}^2\sqrt{(\mathfrak{T} - 1)(\mathfrak{n}^2\mathfrak{T} - (\mathfrak{n} - 2)^2)}} \quad (31)$$

This derivative is positive, so PWL^c decreases with \mathfrak{T} . ■

Proof of Proposition 18: Let

$$\begin{aligned} \alpha &= \mathfrak{p} + \frac{\mathfrak{p}}{\frac{\mathfrak{p}}{\mathfrak{p} - \mathfrak{c}} - \rho(\mathfrak{n} - 1)} \\ \beta &= \frac{\mathfrak{p} \left(\frac{\mathfrak{p}}{\mathfrak{p} - \mathfrak{c}} - \rho(\mathfrak{n} - 2) \right)}{\mathfrak{x}_i \left(\frac{\mathfrak{p}}{\mathfrak{p} - \mathfrak{c}} + \rho \right) \left(\frac{\mathfrak{p}}{\mathfrak{p} - \mathfrak{c}} - \rho(\mathfrak{n} - 1) \right)} \\ \gamma &= \frac{\mathfrak{p}\rho}{\mathfrak{x}_i \left(\frac{\mathfrak{p}}{\mathfrak{p} - \mathfrak{c}} + \rho \right) \left(\frac{\mathfrak{p}}{\mathfrak{p} - \mathfrak{c}} - \rho(\mathfrak{n} - 1) \right)} \end{aligned}$$

It is straightforward to prove that $\alpha > \mathfrak{c}$ and $\beta > 0$, and that $\beta > \gamma \geq 0$ for $\frac{\mathfrak{p}}{\mathfrak{p} - \mathfrak{c}} > \rho(\mathfrak{n} - 1) \geq 0$. One can easily show that the linear market $\{\alpha, \beta, \gamma, \mathfrak{c}, \mathfrak{n}\}$ yields a Bertrand equilibrium where $p_i^b = \mathfrak{p}$ and $x_i^b(\mathfrak{p}, \mathfrak{p}_{-i}) = \mathfrak{x}_i$, and find PWL^b by plugging the values of β and γ in (15). ■

Proof of Proposition 19: In a Bertrand equilibrium $\frac{\mathfrak{p}}{\mathfrak{p} - \mathfrak{c}} = \varepsilon$. From (17), we therefore obtain

$$\begin{aligned} \frac{\partial PWL^b}{\partial \mathfrak{n}} &= -\frac{2\varepsilon\rho(\varepsilon - \rho(\mathfrak{n} - 1))}{(2\varepsilon - \rho(\mathfrak{n} - 1))^3} < 0, \\ \frac{\partial PWL^b}{\partial \varepsilon} &= \frac{2(\mathfrak{n} - 1)\rho(\varepsilon - \rho(\mathfrak{n} - 1))}{(2\varepsilon - \rho(\mathfrak{n} - 1))^3} > 0, \end{aligned}$$

$$\frac{\partial PWL^b}{\partial \rho} = -\frac{2(n-1)\varepsilon(\varepsilon - \rho(n-1))}{(2\varepsilon - \rho(n-1))^3} < 0.$$

From these formulae the proposition follows. ■

Proof of Lemma 20: Social welfare in a Cournot equilibrium, denoted by W^c , is given by

$$W^c = \sum_{i=1}^n (\alpha_i - c_i) x_i^c - \frac{\beta}{2} \sum_{i=1}^n x_i^{c2} - \frac{\gamma}{2} \sum_{i=1}^n x_i^c \sum_{j \neq i} x_j^c = \frac{3\beta - \gamma}{2} H^c x^{c2} + \frac{\gamma}{2} x^{c2}$$

Using the definition of H^o , social welfare in the optimum is

$$W^o = \frac{\beta - \gamma}{2} H^o x^{o2} + \frac{\gamma}{2} x^{o2}.$$

Plugging the values of W^c and W^o into PWL^c yields

$$PWL^c = 1 - \frac{W^c}{W^o} = 1 - \left(\frac{x^c}{x^o} \right)^2 \frac{H^c(3\beta - \gamma) + \gamma}{H^o(\beta - \gamma) + \gamma},$$

while plugging in the values of x^c and x^o yields formula (24). ■

Proof of Lemma 21: Social welfare in a Bertrand equilibrium, denoted by W^b , is given by

$$W^b = \frac{(\beta - \gamma)(3\beta + \gamma(3n - 4))}{2(\beta + \gamma(n - 2))} H^b x^{b2} + \frac{\gamma}{2} x^{b2}.$$

Let PWL^b be the percentage of welfare losses in a Bertrand equilibrium.

$$PWL^b = 1 - \frac{W^b}{W^o} = 1 - \left(\frac{x^b}{x^o} \right)^2 \frac{H^b(\beta - \gamma)(3\beta + \gamma(3n - 4)) + \gamma(\beta + \gamma(n - 2))}{(H^o(\beta - \gamma) + \gamma)(\beta + \gamma(n - 2))}$$

Plugging in the values of x^b and x^o , we obtain the formula above. ■

Proof of Proposition 22: Computing $\frac{\partial PWL^c}{\partial H^c}(m = n)$, we obtain

$$\frac{\frac{\gamma}{\beta} \left(1 - \frac{\gamma}{\beta}\right) \left(1 + \frac{\gamma}{\beta}(n-1)\right) \left(4 + \frac{\gamma}{\beta}(n-2)\right)}{\left(4H^c \left(1 + \frac{\gamma}{\beta}(n-2)\right) + \left(\frac{\gamma}{\beta}\right)^3 (H^c - 1)(n-1) + \left(\frac{\gamma}{\beta}\right)^2 (n-2 + H^c(5-4n))\right)^2},$$

which is positive for $\frac{\gamma}{\beta} > 0$. Also, $\frac{\partial PWL^b}{\partial H^b}(m = n)$ is equal to

$$\frac{\frac{\gamma}{\beta} \left(1 - \frac{\gamma}{\beta}\right) \left(1 + \frac{\gamma}{\beta}(n-1)\right) \left(1 + \frac{\gamma}{\beta}(n-2)\right) \left(4 + 5\frac{\gamma}{\beta}(n-2) + \left(\frac{\gamma}{\beta}\right)^2 (6 + (n-6)n)\right)}{\left(4H^b \left(1 + 2\frac{\gamma}{\beta}(n-2)\right) + \left(\frac{\gamma}{\beta}\right)^3 (3 - H^b(3-2n)^2 + (n-3)n) + \left(\frac{\gamma}{\beta}\right)^2 (n-2 + H^b(21 + 4(n-5)n))\right)^2},$$

which is positive for $\frac{\gamma}{\beta} > 0$. ■

Proof of Lemma 24: The FOC of profit maximization for firm i is

$$r \left(\sum_{i=1}^n x_i^\delta \right)^{\frac{r}{\delta}-1} \delta x_i^{\delta-1} - c = 0. \quad (32)$$

The left-hand side of (32) is decreasing in x_i , so the second-order condition holds. In a symmetric equilibrium where all firms produce the same output, denoted by x_i^* , we have

$$x_i^* = \left(\frac{r\delta}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{1}{1-r}}, \quad p^* = \frac{c}{\delta} \quad \text{and} \quad U^* = n^{\frac{r}{\delta}} \left(\frac{r\delta}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{r}{1-r}}. \quad (33)$$

In this equilibrium, the social welfare is

$$W^* = n^{\frac{r}{\delta}} \left(\frac{r\delta}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{r}{1-r}} - nc \left(\frac{r\delta}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{1}{1-r}}.$$

In the optimum the price equals the marginal cost. Thus, $r \left(\sum_{i=1}^n x_i^\delta \right)^{\frac{r}{\delta}-1} x_i^{\delta-1} = c$. From this we get

$$x_i^o = \left(\frac{r}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{1}{1-r}} \quad \text{and} \quad W^o = n^{\frac{r}{\delta}} \left(\frac{r}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{r}{1-r}} - nc \left(\frac{r}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{1}{1-r}}, \quad (34)$$

where x_i^o and W^o stand for the output and social welfare in the optimum. W^o is increasing in n , so in the full optimum the planner would choose a number of firms equal to n .

Consequently, the percentage of welfare losses is:

$$PWL^s = 1 - \frac{W^*}{W^o} = 1 - \frac{n^{\frac{r}{\delta}} \left(\frac{r\delta}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{r}{1-r}} - nc \left(\frac{r\delta}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{1}{1-r}}}{n^{\frac{r}{\delta}} \left(\frac{r}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{r}{1-r}} - nc \left(\frac{r}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{1}{1-r}}} = 1 - \delta^{\frac{1}{1-r}} \frac{\frac{1}{\delta} - r}{1 - r}. \blacksquare$$

Proof of Proposition 26: Let δ and r be such that

$$\left(\frac{r\delta}{cn^{1-\frac{r}{\delta}}} \right)^{\frac{1}{1-r}} = \mathfrak{x}_i \quad \text{and} \quad \frac{c}{\delta} = \mathfrak{p}$$

The previous equations yield

$$\delta = \frac{c}{p} \quad \text{and} \quad r = \frac{ProductLog(\mathfrak{n} p \mathfrak{x}_i \left(\frac{p}{c} \ln \mathfrak{n} + \ln \mathfrak{x}_i \right))}{\frac{p}{c} \ln \mathfrak{n} + \ln \mathfrak{x}_i}$$

It is straightforward to check that $0 < \delta < 1$ and $0 < r < 1$ (using the condition $\frac{c}{p} < \frac{\ln \mathfrak{n}}{\ln \mathfrak{n} + \ln p}$). Then by construction the CES market $\{\delta, r, c, \mathfrak{n}\}$ yields an equilibrium where $p^* = p$ and $x_i^* = \mathfrak{x}_i$. Plugging δ and r into (29), we get the formula for PWL^s as a function of an observation $\{p, \mathfrak{x}_i, c, \mathfrak{n}\}$. \blacksquare

CHAPTER IV

WELFARE LOSSES IN MODELS OF HORIZONTAL AND VERTICAL DIFFERENTIATION

4.1 *Introduction*

One of the main issues in the theory of Industrial Organization is the inefficiencies yielded by the decentralized nature of equilibrium. In a pathbreaking contribution, Harberger (1954) provided a quantitative estimate of these inefficiencies. His paper, no matter how debatable, has generated an enormous amount of literature trying to prove or disprove his main findings. Unfortunately, much of this literature did not pay attention to the subtleties implied by the theory of Industrial Organization. Recently, theoretical background to approach this problem has been provided by Cable et al. (1994), McHardy (2000), Anderson and Renault (2003), Johari and Tsitsiklis (2005), Corchón (2008) and Corchón and Zudenkova (2009).

Another strand of literature is Algorithmic Game Theory. This is a new scientific branch arising from the confluence of Computer Science, with its emphasis on computing algorithms, and Game Theory with its emphasis on rational players. For a good introduction see Nisan et al. (2007). The theory has several goals. Of particular interest to us is the goal of quantifying the extent of inefficiencies yielded by the non-cooperative equilibrium.¹ This is done by considering the ratio between the values of the objective function representing social values in the equilibrium and in the optimum.²

Our paper aims to contribute to the convergence of these two strands of literature by studying the percentage of welfare losses in models of horizontal (Hotelling 1929—amended by d’Aspremont, Gabszewicz and Thisse 1979, and Economides 1984—and Salop 1979) and vertical differentiation (Gabszewicz and Thisse 1979, 1980, and Shaked and Sutton 1982,

¹Other goals are to explain how equilibrium is reached and the design of resource allocation mechanisms with good properties, with special attention to computational problems. See Nisan et al. (2007).

²The measure considered by Harberger (H) and that used in algorithmic game theory (A) are related as follows: $H = 1 - A$.

1983). In all these models firms act strategically. The motivation for our study is that these models are very different from those in which welfare losses have been investigated. In the models considered here, there is a continuum of consumers (and not a representative consumer) and each consumer consumes, at most, one unit of the differentiated good (and not many units). Also, in the models of Hotelling and Shaked and Sutton firms are free to locate in the product space. In both respects these models are closer to the model of a communication network by Acemoglu, Bimpikisb and Ozdaglar (2009).³ As we will see, these differences make the analysis of welfare losses here very different from those analyzed before.

Section 4.2 considers the Hotelling model. Like Harberger (1954), we consider the percentage of welfare losses (PWL) defined as the percentage at which equilibrium welfare falls short of the optimum. We first study how PWL depends on the basic parameters that define the market such as the reservation price (α), transportation cost (β) and marginal costs (c). We find that PWL depends on $\frac{\alpha-c}{\beta}$ but not in a monotonic way. The reason is that a change in $\frac{\alpha-c}{\beta}$ not only changes welfare for given locations but it causes firms to reallocate. Next, we study if PWL can be recovered from observable variables such as prices, marginal costs, the distance between brands and the percentage of market coverage. We find that, in most cases, PWL can be calculated from location and market coverage alone. PWL decreases with market coverage (unless in the optimum the whole market should not be covered, in which case PWL is constant) and with the distance from the market edges as it goes to the optimal location. Prices, marginal costs (and thus markups) and demand elasticities do not help to find PWL . Finally, PWL might be large despite price competition. This shows that misallocation arising from the wrong location could be very significant, especially when not all the market is covered.

Section 4.3 considers a model of a circular city (Salop 1979). Here PWL depends on $\frac{\alpha-c}{\beta}$ but also on the number of firms (n) and the form of the transportation cost (γ) which in Hotelling was quadratic ($\gamma = 2$) and here is either linear ($\gamma = 1$) or quadratic ($\gamma = 2$). As

³But the models considered here and that of Acemoglu, Bimpikisb and Ozdaglar are different too. For instance, in the latter, equilibrium is not necessarily unique.

in the Hotelling model, PWL decreases with the market coverage, unless in the optimum the whole market should not be covered, in which case PWL is constant. Here, PWL can be calculated in all cases. In other words, the indeterminacy that occurred in the Hotelling model does not arise here. As in the Hotelling model and for the same reason, PWL is independent of demand elasticities and markups. Also, since here there are no misallocations due to firms being in the wrong locations, PWL tends to be smaller than in the Hotelling model. But these losses may be large, up to 25%. Finally, PWL is not monotonic in γ for given market coverage. The reason is that a change in γ changes welfare both in equilibrium and in the optimal allocations.

In Section 4.4 we study oligopolistic competition under quality differentiation, see Gabszewicz and Thisse (1979, 1980) and Shaked and Sutton (1982, 1983). We assume that the parameter that measures the taste for quality is uniformly distributed across the population of consumers, with extrema a and b . Here we have two kinds of equilibria: those in which the whole market is covered and those in which not all the market is covered. We find that PWL is a *discontinuous* function of $\frac{b}{a}$ alone with a maximum of about 8.33%. The discontinuity is caused by the fact that when $\frac{b}{a} = 8$, firm 1 freezes the quality of its good and does not serve consumers with low taste for quality. Thus the discontinuity arises at the point in which the market becomes uncovered.

We show that under quality differentiation welfare losses can be found from the relative prices and the degree of market coverage. When the whole market is covered, PWL is single-peaked in relative prices reaching a maximum at an interior point. Thus an increase in relative prices can decrease or increase relative welfare losses. When not all the market is covered, PWL depends only on market coverage, in a decreasing way, as expected.

Section 4.5 sums up our findings.

4.2 *Horizontal Differentiation: The Hotelling Model*

There are two firms producing a differentiated good. Consumers purchase either one unit or none of the differentiated good according to preferences, prices and the distribution of the two brands in product space. Brands are located in the interval $[0, 1]$. Each consumer

has a most-preferred brand specification τ . Consumers are uniformly distributed along $[0, 1]$ with density 1. A brand located at point x_i , $i = 1, 2$, is valued for the consumer at point τ according to $U(x_i, \tau) = \alpha - \beta d^\gamma$ where α stands for the reservation price, $d = |x_i - \tau|$ is the Euclidean distance between x_i and τ , β and γ measure the importance of transportation costs, and α and β are positive. The decision rule of consumer τ is: purchase one unit of brand x_i if $\max_i [U(x_i, \tau) - p_i] \geq 0$, where p_i is the price of the brand x_i , $i = 1, 2$. The marginal cost of production is $c < \alpha$.

The model where $\gamma = 1$ is not easily tractable since profit functions are discontinuous and nonconcave.⁴ To overcome these difficulties we assume that $\gamma = 2$. Summing up:

Definition 28 *A Linear Horizontal Market is a list of positive real numbers $\{\alpha, \beta, c\}$ with $\alpha > c$.*⁵

Let us consider a two-stage game. In the first stage, firms choose their locations x_1 and x_2 simultaneously. In the second stage, they choose prices simultaneously. Without loss of generality, assume that $x_2 \geq x_1$. Firm i 's profit is $\pi_i = (p_i - c) D_i$ where D_i is the demand of firm i . It is easy to show that profit functions are continuous and concave and a Subgame Perfect Nash Equilibrium exists.⁶

We consider three symmetric equilibrium configurations: local monopolistic equilibrium, kinked equilibrium and competitive equilibrium.⁷ We characterize equilibria where consumers at the edges of the market buy the differentiated good.⁸

Local monopolistic equilibrium. At this equilibrium, some consumers lying between two firms do not purchase the differentiated good, so the market is not fully covered. Each

⁴For further details see d'Aspremont, Gabszewicz and Thisse (1979) and Economides (1984).

⁵"Linear" relates to the linear form of the product space. "Horizontal" refers to the form of product differentiation. In Section 4.4 we will speak of a "Linear Vertical" market in which product differentiation is vertical.

⁶See Economides (1986) for the general case of $\gamma \leq 2$. Economides (1986) showed existence for $\gamma \in (1.26, 2]$.

⁷Salop (1979) used the term "kinked" for an equilibrium where the markets just touch and there is no tangency of demand. Economides (1984) used the term "touching" for such an equilibrium.

⁸Economides (1984) studied the case of "not-too-high" reservation price where consumers at the edges of the market prefer not to purchase the differentiated good. He showed that under linear transportation cost function the equilibrium of the locations game is a local monopolistic one. The reason is that in the "competitive region" firms have incentives to relocate marginally away from each other and reach the "kinked" region. While in the "kinked" region firms still want to relocate away from each other and reach the "local monopolistic" region.

firm charges monopoly price p_m . A consumer with preferred brand $\tau \in (x_1, \frac{1}{2})$ is indifferent between purchasing from firm 1 and not purchasing the differentiated good if $\alpha - \beta(\tau - x_1)^2 - p_m = 0$. Thus, firm 1's demand is $D_1(p_m) = \tau = x_1 + \sqrt{\frac{\alpha - p_m}{\beta}}$. In the second stage firm 1's profit maximization with respect to p_m yields

$$p_m(x_1) = \frac{1}{9} \left(6\alpha + 3c - 2\beta x_1^2 + 2\sqrt{\beta x_1^2 (3(\alpha - c) + \beta x_1^2)} \right).$$

Plugging $p_m(x_1)$ into firm 1's profit yields

$$\pi_1(x_1) = (p_m(x_1) - c) \left(x_1 + \sqrt{\frac{\alpha - p_m(x_1)}{\beta}} \right).$$

After tedious calculations one finds that $\frac{\partial \pi_1}{\partial x_1} > 0$, so firms have incentives to relocate towards the market center still maintaining local monopoly power. Firms will move to the market center until consumers at the edges of the market are just indifferent between buying the differentiated good and not. One can check that in this case the firms' marginal relocation tendency becomes zero.

Consumers at the edges of the market are indifferent between buying the good and not buying it, which amounts to $x_1 - \sqrt{\frac{\alpha - p_m(x_1)}{\beta}} = 0$. Consumers at the market center do not buy the differentiated good, which amounts to $x_1 + \sqrt{\frac{\alpha - p_m(x_1)}{\beta}} < \frac{1}{2}$. These two conditions yield a local monopolistic equilibrium: firm 1 chooses $x_1^* = \sqrt{\frac{\alpha - c}{5\beta}}$, and by symmetry firm 2 chooses $x_2^* = 1 - \sqrt{\frac{\alpha - c}{5\beta}}$. Both firms charge the same price $p_m^* = \frac{4\alpha + c}{5}$. This equilibrium exists for $\frac{\alpha - c}{\beta} < \frac{5}{16}$. Thus, a Local Monopolistic Equilibrium for a linear horizontal market $\{\alpha, \beta, c\}$ with $\frac{\alpha - c}{\beta} < \frac{5}{16}$ is a price $p_m^* = \frac{4\alpha + c}{5}$ and brand locations $\{x_{m1}^*, x_{m2}^*\} = \left\{ \sqrt{\frac{\alpha - c}{5\beta}}, 1 - \sqrt{\frac{\alpha - c}{5\beta}} \right\}$.

Kinked equilibrium. At a kinked equilibrium, markets just "touch". A consumer with preferred brand specification $\tau = \frac{1}{2}$ is indifferent between purchasing from firm 1 or from firm 2 at price p_k and not purchasing the differentiated good if $\alpha - \beta\left(\frac{1}{2} - x_1\right)^2 - p_k = 0$. Thus, $p_k(x_1) = \alpha - \beta\left(\frac{1}{2} - x_1\right)^2$. At the same time, firms still enjoy local monopolistic power, therefore $p_k(x_1) = p_m(x_1)$, which yields

$$x_1^* = 1 - \frac{1}{2} \sqrt{4\frac{\alpha - c}{\beta} + 1} \text{ and } p_k^* = \frac{1}{2} \left(2c - \beta + \sqrt{\beta(4(\alpha - c) + \beta)} \right).$$

In the kinked equilibrium firms behave as local monopolists but maintain full market coverage so $x_1^* > 0$ or $\frac{\alpha-c}{\beta} < \frac{3}{4}$. The consumers at the edges of the market purchase the differentiated good so $\alpha - \beta (x_1^*)^2 - p_k^* \geq 0$, which simplifies to $\frac{\alpha-c}{\beta} \geq \frac{5}{16}$.

Note that for $\frac{3}{4} \leq \frac{\alpha-c}{\beta} < \frac{5}{4}$ there is a kinked equilibrium with $\{x_1^*, x_2^*\} = \{0, 1\}$ and $p_k^* = \alpha - \frac{\beta}{4}$, which is an intermediate case between the kinked equilibrium described above and the competitive equilibrium, which is analyzed below. Therefore, a Kinked Equilibrium for a linear horizontal market $\{\alpha, \beta, c\}$ with $\frac{5}{16} \leq \frac{\alpha-c}{\beta} < \frac{3}{4}$ is a price $p_k^* = \frac{1}{2} \left(2c - \beta + \sqrt{\beta(4(\alpha-c) + \beta)} \right)$ and brand locations $\{x_{k1}^*, x_{k2}^*\} = \left\{ 1 - \frac{1}{2} \sqrt{4\frac{\alpha-c}{\beta} + 1}, \frac{1}{2} \sqrt{4\frac{\alpha-c}{\beta} + 1} \right\}$. Moreover, a Kinked Equilibrium for a linear horizontal market $\{\alpha, \beta, c\}$ with $\frac{3}{4} \leq \frac{\alpha-c}{\beta} < \frac{5}{4}$ is a price $p_k^* = \alpha - \frac{\beta}{4}$ and brand locations $\{x_{k1}^*, x_{k2}^*\} = \{0, 1\}$.

Competitive equilibrium. A consumer with preferred brand specification $\tau \in (x_1, x_2)$, is indifferent between purchasing brand x_1 and purchasing brand x_2 if

$$\alpha - \beta (\tau - x_1)^2 - p_1 = \alpha - \beta (x_2 - \tau)^2 - p_2 \Rightarrow \tau = \frac{p_2 - p_1}{2\beta(x_2 - x_1)} + \frac{x_1 + x_2}{2},$$

so the demands $D_1(p_1, p_2)$ and $D_2(p_1, p_2)$ faced by firms 1 and 2 read

$$\begin{aligned} D_1(p_1, p_2) &\equiv \tau = \frac{p_2 - p_1}{2\beta(x_2 - x_1)} + \frac{x_1 + x_2}{2} \\ D_2(p_1, p_2) &\equiv 1 - \tau = 1 - \frac{p_2 - p_1}{2\beta(x_2 - x_1)} - \frac{x_1 + x_2}{2} \end{aligned}$$

Firm i 's profit maximization with respect to p_i yields

$$p_1 = c + \frac{\beta}{3} (x_2 - x_1) (2 + x_1 + x_2) \text{ and } p_2 = c + \frac{\beta}{3} (x_2 - x_1) (4 - x_1 - x_2),$$

and corresponding profits become

$$\pi_1 = \frac{\beta}{18} (x_2 - x_1) (2 + x_1 + x_2)^2 \text{ and } \pi_2 = \frac{\beta}{18} (x_2 - x_1) (4 - x_1 - x_2)^2.$$

The "marginal relocation tendency of firms" reads $\frac{\partial \pi_1}{\partial x_1} < 0$ and $\frac{\partial \pi_2}{\partial x_2} > 0$. Thus, the firms have incentives to relocate marginally away from each other. The equilibrium has two firms locating at the two extremes of the product space (maximal differentiation) $x_1^* = 0$ and $x_2^* = 1$ and charging the same price $p_c^* = c + \beta$. At the competitive equilibrium the entire market is covered, so $\alpha - \beta \left(\frac{1}{2}\right)^2 - p_c^* \geq 0$, which amounts to $\frac{\alpha-c}{\beta} \geq \frac{5}{4}$. Finally, a Competitive

Equilibrium for a linear horizontal market $\{\alpha, \beta, c\}$ with $\frac{\alpha-c}{\beta} \geq \frac{5}{4}$ is a price $p_c^* = c + \beta$ and a list of brand locations $\{x_{c1}^*, x_{c2}^*\} = \{0, 1\}$.

Thus, the equilibrium configuration depends on the values taken by $\frac{\alpha-c}{\beta}$.

- For low values of $\frac{\alpha-c}{\beta}$ (i.e. $\frac{\alpha-c}{\beta} < \frac{5}{16}$), there exists a local monopolistic symmetric equilibrium where the firms enjoy monopolistic power at the market edges while consumers in the centre of the market do not purchase the differentiated commodity. The higher the value of $\frac{\alpha-c}{\beta}$, the closer firm 1 (resp. firm 2) to location $\frac{1}{4}$ (resp. $\frac{3}{4}$).

- If $\frac{\alpha-c}{\beta} = \frac{5}{16}$ there exists a kinked equilibrium where the whole market is covered and $\{x_{k1}^*, x_{k2}^*\} = \{\frac{1}{4}, \frac{3}{4}\}$. In this equilibrium firms enjoy all the monopolistic power: they sell to consumers both at the market edges and in the market center, and still do not become involved in the competition between each other.

- If $\frac{5}{16} < \frac{\alpha-c}{\beta} < \frac{3}{4}$ there is a kinked equilibrium where the whole market is covered, firms do not compete for the market center consumers (the markets just touch) but do not extract all the possible surplus from the market edges consumers: the consumer at the edges of the market get positive surplus by purchasing the differentiated commodity.

- If $\frac{3}{4} \leq \frac{\alpha-c}{\beta} < \frac{5}{4}$ there exists a kinked equilibrium where the whole market is covered, firms are situated at the edges of the market $\{x_{k1}^*, x_{k2}^*\} = \{0, 1\}$ and do not compete for the market center consumer (the markets just touch).

- Finally, for high values of $\frac{\alpha-c}{\beta}$ (i.e. $\frac{\alpha-c}{\beta} \geq \frac{5}{4}$) there is a competitive equilibrium where the entire market is covered, the firms are situated at the edges of the market $\{x_{c1}^*, x_{c2}^*\} = \{0, 1\}$ and compete for the market center consumers.

Define social welfare W as the gross consumers' surplus minus costs (i.e., the marginal cost and the consumers' transportation costs):

$$W = 2 \int_0^{\tau} \left(\alpha - c - \beta (t - x_1^*)^2 \right) dt$$

which is equal in equilibrium to

$$W^* = \begin{cases} \frac{56}{15\sqrt{5}} (\alpha - c) \sqrt{\frac{\alpha-c}{\beta}} & \text{if } \frac{\alpha-c}{\beta} < \frac{5}{16} \\ \frac{3}{4} \beta \sqrt{4 \frac{\alpha-c}{\beta}} + 1 - \frac{5}{6} \beta & \text{if } \frac{5}{16} \leq \frac{\alpha-c}{\beta} < \frac{3}{4} \\ \alpha - c - \frac{\beta}{12} & \text{if } \frac{\alpha-c}{\beta} \geq \frac{3}{4} \end{cases}$$

Note that the social welfare does not depend on the equilibrium price, which is just a transfer from the consumers to the producers. In our framework the social welfare is only affected by the market coverage and firms' locations, which determine the consumers' transportation costs and the consumers' surplus net of marginal cost. This is the reason why the social welfare in a kinked equilibrium for $\frac{3}{4} \leq \frac{\alpha-c}{\beta} < \frac{5}{4}$ is the same as the social welfare in a competitive equilibrium for $\frac{\alpha-c}{\beta} \geq \frac{5}{4}$. Indeed, in both cases the whole market is covered and the firms are located at the market edges.

A social planner would choose the price equal to marginal cost and the firms' locations to maximize the social welfare. One can check that the social welfare in the optimum, denoted by W^o , is equal to

$$W^o = \begin{cases} \frac{8}{3}(\alpha - c) \sqrt{\frac{\alpha-c}{\beta}} & \text{if } \frac{\alpha-c}{\beta} < \frac{1}{16} \\ \alpha - c - \frac{\beta}{48} & \text{if } \frac{\alpha-c}{\beta} \geq \frac{1}{16} \end{cases}$$

where the first line corresponds to the case where not all the market is covered and consumers at the market center do not purchase the differentiated commodity; the second line corresponds to the case where the entire market is covered and the firms are located at $\{x_1^o, x_2^o\} = \{\frac{1}{4}, \frac{3}{4}\}$.

The percentage of welfare losses is defined as

$$PWL \equiv 1 - \frac{W^*}{W^o}$$

and is equal to

$$PWL = \begin{cases} 1 - \frac{7}{5\sqrt{5}} \approx 0.3739 & \text{if } \frac{\alpha-c}{\beta} < \frac{1}{16} \\ 1 - \frac{\frac{56}{15\sqrt{5}} \sqrt{\frac{\alpha-c}{\beta}}}{1 - \frac{1}{48} \frac{\beta}{\alpha-c}} & \text{if } \frac{1}{16} \leq \frac{\alpha-c}{\beta} < \frac{5}{16} \\ 1 - \frac{\frac{3}{4} \sqrt{4 \frac{\alpha-c}{\beta} + 1} - \frac{5}{6}}{\frac{\alpha-c}{\beta} - \frac{1}{48}} & \text{if } \frac{5}{16} \leq \frac{\alpha-c}{\beta} < \frac{3}{4} \\ 1 - \frac{\frac{\alpha-c}{\beta} - \frac{1}{12}}{\frac{\alpha-c}{\beta} - \frac{1}{48}} & \text{if } \frac{\alpha-c}{\beta} \geq \frac{3}{4} \end{cases} \quad (35)$$

Figure 11 depicts PWL as a function of $\frac{\alpha-c}{\beta}$. Note that PWL is not monotonic in $\frac{\alpha-c}{\beta}$. The reason is that change in $\frac{\alpha-c}{\beta}$ not only changes welfare for given locations but it causes reallocation effects that may overcome the latter effect. For instance, a decrease in transportation costs β makes the economy more competitive, but at the same time causes firms to relocate away from each other increasing monopoly power.

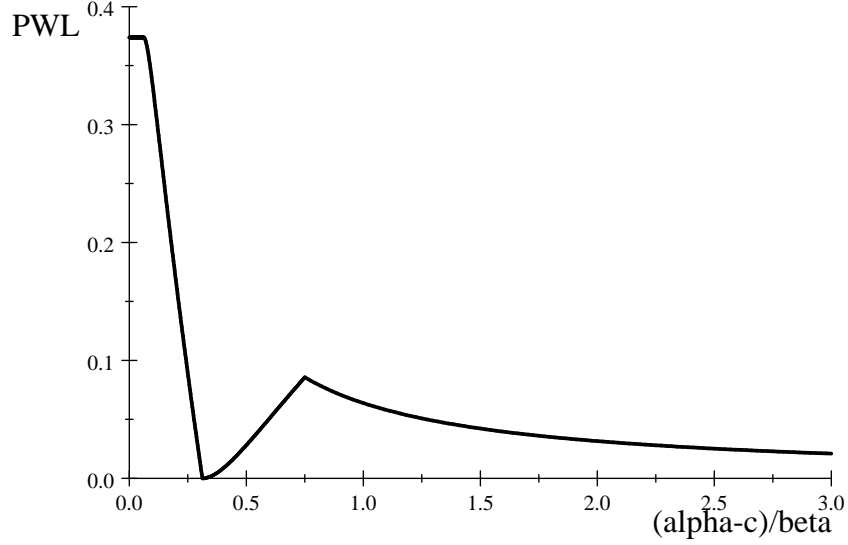


Figure 11: PWL as a function of $\frac{\alpha-c}{\beta}$.

- For $\frac{\alpha-c}{\beta} < \frac{1}{16}$ PWL is constant. In this case the market is not covered either in equilibrium or in the optimum. Welfare losses are due to the firms' monopolistic behavior.
- For $\frac{1}{16} \leq \frac{\alpha-c}{\beta} < \frac{5}{16}$, PWL is decreasing in $\frac{\alpha-c}{\beta}$. In this case in equilibrium the market is not covered, while in the optimum it should be covered. Indeed, with increase of $\frac{\alpha-c}{\beta}$ the equilibrium configuration becomes closer to the optimum configuration $\{x_1^o, x_2^o\} = \{\frac{1}{4}, \frac{3}{4}\}$.
- For $\frac{\alpha-c}{\beta} = \frac{5}{16}$ $PWL = 0$ since equilibrium configuration is the same as the optimum one: the whole market is covered and the firms are located at $\{\frac{1}{4}, \frac{3}{4}\}$.
- When $\frac{5}{16} < \frac{\alpha-c}{\beta} < \frac{3}{4}$ PWL is increasing in $\frac{\alpha-c}{\beta}$. In this case the higher $\frac{\alpha-c}{\beta}$ is, the closer the firms locate to the edges of the market (so the higher the consumers' transportation costs are). This effect exacerbates welfare losses.
- Finally, when $\frac{\alpha-c}{\beta} \geq \frac{3}{4}$ PWL is decreasing in $\frac{\alpha-c}{\beta}$. Here in equilibrium, the firms locate at the market edges and compete for the consumers located at the market center. Thus, PWL is decreasing in $\frac{\alpha-c}{\beta}$.

So far we have analyzed the relationship between PWL and the parameters defining a linear horizontal market $\{\alpha, \beta, c\}$. Let us now relate PWL with observable variables. An *Observation* is a tuple $\{p, c, r, m\}$ of market price p , marginal cost c ($p > c$), the relative distance from the market edges to brand locations $r \in [0, \frac{1}{4}]$ and the percentage of market

coverage $\mathbf{m} \in [0, 1]$. We assume that the marginal cost is observable because under constant returns, the marginal cost equals the average variable cost, which in principle can be observed (wages, raw materials, etc.). Then, we have the following:

Proposition 29 *Given an observation $\{\mathbf{p}, \mathbf{c}, \mathbf{r}, \mathbf{m}\}$ there is a linear horizontal market $\{\alpha, \beta, \mathbf{c}\}$ such that $\{\mathbf{p}, \mathbf{r}, 1 - \mathbf{r}\}$ is*

- i) a local monopolistic equilibrium for this market when $\mathbf{m} < 1$ and $0 \leq \mathbf{r} < \frac{1}{4}$;*
- ii) a kinked equilibrium for this market when $\mathbf{m} = 1$ and $0 < \mathbf{r} \leq \frac{1}{4}$;*
- iii) either a kinked equilibrium or a competitive equilibrium for this market when $\mathbf{m} = 1$ and $\mathbf{r} = 0$;*

PWL in each case is given by

$$PWL = \begin{cases} 1 - \frac{7}{5\sqrt{5}} \approx 0.3739 & \text{if } \mathbf{m} < 1, 0 \leq \mathbf{r} < \frac{1}{4\sqrt{5}} \\ 1 - \frac{896\mathbf{r}^3}{240\mathbf{r}^2 - 1} & \text{if } \mathbf{m} < 1, \frac{1}{4\sqrt{5}} \leq \mathbf{r} < \frac{1}{4} \\ 1 - \frac{32 - 72\mathbf{r}}{35 - 48\mathbf{r}(2 - \mathbf{r})} & \text{if } \mathbf{m} = 1, 0 < \mathbf{r} \leq \frac{1}{4} \\ \in (0, \frac{3}{35}] , \frac{3}{35} \approx 0.0857 & \text{if } \mathbf{m} = 1, \mathbf{r} = 0 \end{cases} \quad (36)$$

Proofs of this and other propositions are given in the Appendix.

Figure 12 depicts PWL as a function of observables. In general, as expected, PWL decreases with market coverage (unless in the optimum the whole market should not be covered, in which case PWL is constant) and with the distance \mathbf{r} from the market edges as it goes to the optimal location $\frac{1}{4}$. Other points are worth discussion.

Firstly, PWL can be calculated from location and market coverage in three out of four cases in (36). Only in the case where the whole market is covered and the firms locate at the market edges, i.e. $\mathbf{m} = 1$ and $\mathbf{r} = 0$, could PWL be any number between zero and $\frac{3}{35}$, even if price and marginal cost are observed.⁹ Knowledge of demand elasticity, denoted by ε , cannot be used to break the indeterminacy of PWL . In the case of a kinked equilibrium, the demand function is not differentiable so demand elasticity is not well defined. In the case of a competitive equilibrium, from the first order condition of profit maximization

⁹ PWL could be calculated if the reservation price is observed. The latter is usually thought to be private information but, in some cases, can be elicited by the mechanism of Becker, DeGroot and Marschak (1964). For the limitations of this mechanism see Horowitz (2006) and the references there.

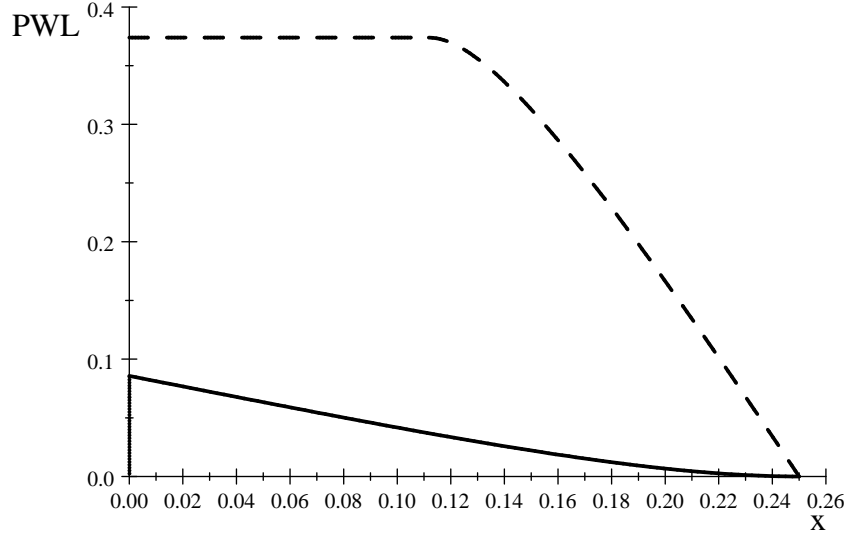


Figure 12: PWL as a function of observable relative distance \mathfrak{x} from the market edges to brand locations when the market is not covered, $\mathfrak{m} < 1$ (dash), and when the market is covered, $\mathfrak{m} = 1$ (solid).

$\varepsilon = \frac{p}{p-c}$, so knowledge of ε is redundant. The same argument applies if the cross elasticity of demand $\frac{\partial D_i}{\partial p_j} \frac{p_j}{D_i}$, denoted by ρ , is observable since in our model, in equilibrium, $\rho = \varepsilon$.¹⁰

Secondly, PWL is independent from demand elasticities (own and cross) and markups. This is explained by the fact that as demand is totally inelastic, a high price, unless it induces not to buy the good, does not cause welfare losses. As we remarked before, an increase in price just redistributes the surplus between consumers and firms. This makes a difference with models in which consumers may buy several goods where demand elasticities and markups can be used to find PWL even though their impact is sometimes counterintuitive. See Corchón and Zudenkova (2009).

Thirdly, PWL might be large, larger than in the Cournot model with linear demand and cost functions, which in the duopoly case is around 11% (Anderson and Renault 2003). And this occurs despite price competition in the Hotelling model. This shows that misallocation arising from the wrong location could be very significant, especially when not all the market

¹⁰ However, if it could be determined if the demand function is differentiable, we would know if the market is in a kinked equilibrium—where PWL lies between $\frac{3}{59}$ and $\frac{3}{35}$ —or in a competitive equilibrium where PWL lies between 0 and $\frac{3}{59}$.

is covered. See Figure 12.

4.3 Horizontal Differentiation: The Salop Model

Consider the economy described in the previous section with the following changes. Firstly, the product space of the monopolistically competitive industry is a circle with a perimeter equal to 1. Secondly, there are n brands of the differentiated good available at prices p_1, \dots, p_n . As before, each firm is allowed to produce just one brand. Thirdly, firms do not choose their brand location, but are automatically located equidistant from one another on the circle.¹¹ This simplification allows this model to be solved for linear ($\gamma = 1$) and quadratic ($\gamma = 2$) transportation costs. Summing up:

Definition 30 A Circular Market is a list $\{\alpha, \beta, c, n, \gamma\}$, where $\alpha, \beta, c \in \mathbb{R}^+$, $\gamma \in \{1, 2\}$, $n \in \mathbb{N}$, and $\alpha > c$.

Firm i 's profit is $\pi_i = (p_i - c) D_i$ where D_i is the demand firm i faces and p_i is the price chosen by firm i . As before, we consider three symmetric equilibrium configurations: local monopolistic equilibrium, kinked equilibrium and competitive equilibrium.

Local monopolistic equilibrium. At the local monopolistic equilibrium, some consumers lying between two neighboring firms do not purchase the differentiated commodity, so the market is not covered. Each firm charges monopoly price p_m . A consumer with preferred brand specification located at the distance $\tau \in (0, \frac{1}{2n})$ from firm i 's brand specification, is indifferent between purchasing from firm i and not purchasing the differentiated commodity if $\alpha - \beta\tau^\gamma - p_m = 0$. Thus, firm i 's demand is

$$D_i(p_m) = 2\tau = 2 \left(\frac{\alpha - p_m}{\beta} \right)^{\frac{1}{\gamma}}.$$

Firm i 's profit maximization yields $p_m = \frac{c + \gamma\alpha}{1 + \gamma}$. In local monopolistic equilibrium not all the market is covered, which amounts to $\frac{(\alpha - c)n^\gamma}{\beta} < \frac{1 + \gamma}{2^\gamma}$. Thus, a Local Monopolistic Equilibrium for a circular market $\{\alpha, \beta, c, n, \gamma\}$ with $\frac{(\alpha - c)n^\gamma}{\beta} < \frac{1 + \gamma}{2^\gamma}$ is a price $p_m^* = \frac{c + \gamma\alpha}{1 + \gamma}$ and a quantity $D_i(p_m^*) = 2 \left(\frac{\alpha - c}{(1 + \gamma)\beta} \right)^{\frac{1}{\gamma}}$, $i = 1, 2, \dots, n$.

¹¹See Economides (1989) where this assumption emerges in equilibrium in a model where firms decide on locations.

Kinked equilibrium. At a kinked equilibrium, markets just "touch". A consumer with preferred brand specification located at the distance $\tau = \frac{1}{2n}$ from a firm's brand specification, is indifferent between purchasing from a firm or from its closest neighbor at price p_k and not purchasing the differentiated commodity if $\alpha - \beta \left(\frac{1}{2n}\right)^\gamma - p_k = 0$. Thus, $p_k = \alpha - \beta \left(\frac{1}{2n}\right)^\gamma$. In kinked equilibrium the entire market is covered, $D_i = \frac{1}{n}$, but there is no tangency of demand. These conditions amount to $\frac{1+\gamma}{2^\gamma} \leq \frac{(\alpha-c)n^\gamma}{\beta} \leq \frac{1}{2^\gamma} + 1$. Therefore, a Kinked Equilibrium for a circular market $\{\alpha, \beta, c, n, \gamma\}$ with $\frac{1+\gamma}{2^\gamma} \leq \frac{(\alpha-c)n^\gamma}{\beta} \leq \frac{1}{2^\gamma} + 1$ is a price $p_k^* = \alpha - \beta \left(\frac{1}{2n}\right)^\gamma$ and a quantity $D_i(p_k^*) = \frac{1}{n}$, $i = 1, 2, \dots, n$.

Competitive equilibrium. Firms are located equidistant from one another and compete in prices given these locations. Since they are located symmetrically, they will charge the same price p_c in the equilibrium. Firm i has two potential competitors, namely firms $i - 1$ and $i + 1$. Suppose that it chooses price $p_i \equiv p$. A consumer with preferred brand specification located at the distance $\tau \in (0, \frac{1}{n})$ from firm i 's brand specification, is indifferent between purchasing from firm i and from i 's closest neighbor if $\alpha - \beta\tau^\gamma - p = \alpha - \beta \left(\frac{1}{n} - \tau\right)^\gamma - p_c$. Thus, i 's demand reads

$$D_i(p, p_c) = 2\tau = \frac{1}{n} + \frac{p_c - p}{\beta} n^{\gamma-1}.$$

Firm i 's profit maximization yields (in equilibrium $p = p_c$) $p_c = c + \frac{\beta}{n^\gamma}$. In competitive equilibrium all consumers receive positive net surplus so the entire market is covered, which amounts to $\frac{(\alpha-c)n^\gamma}{\beta} > \frac{1}{2^\gamma} + 1$. Finally, a Competitive Equilibrium for a circular market $\{\alpha, \beta, c, n, \gamma\}$ with $\frac{(\alpha-c)n^\gamma}{\beta} > \frac{1}{2^\gamma} + 1$ is a price $p_c^* = c + \frac{\beta}{n^\gamma}$ and a quantity $D_i(p_c^*, p_c^*) = \frac{1}{n}$, $i = 1, 2, \dots, n$.

As before, the equilibrium configuration depends on the underlying parameters. When $\frac{\alpha-c}{\beta}$ and n are small, the market is small (either because the reservation price and/or the number of firms are small or because marginal costs and/or transportation costs β are large) and local monopolies arise.¹² For intermediate values of $\frac{(\alpha-c)n^\gamma}{\beta}$ markets touch and a kinked equilibrium arises. Finally, when $\frac{(\alpha-c)n^\gamma}{\beta}$ is large enough the economy becomes competitive.

¹²Notice that γ appears in both sides of the inequalities defining different equilibria and its effect on them is not straightforward.

Social welfare, W , defined as before is

$$W = 2n \int_0^\tau (\alpha - c - \beta t^\gamma) dt,$$

which in the equilibrium reads for $\gamma = 1, 2$

$$W^* = \begin{cases} \frac{2\gamma(2+\gamma)}{(1+\gamma)^{2+\frac{1}{\gamma}}} \frac{(\alpha-c)^{1+\frac{1}{\gamma}} n}{\beta^{\frac{1}{\gamma}}} & \text{if } \frac{(\alpha-c)n^\gamma}{\beta} < \frac{1+\gamma}{2^\gamma} \\ \alpha - c - \frac{\beta}{2^\gamma(1+\gamma)n^\gamma} & \text{if } \frac{(\alpha-c)n^\gamma}{\beta} \geq \frac{1+\gamma}{2^\gamma} \end{cases}$$

Notice that, again, the social welfare does not depend on the equilibrium price, which is just a transfer from the consumers to the producers. In our framework the social welfare is only affected by the market coverage, which determines the consumers' transportation costs and the consumers' surplus net of marginal cost.

A social planner would choose the price equal to marginal cost, which affects the market coverage and, therefore, the consumers' transportation costs and the consumers' surplus net of marginal cost. It is straightforward to show that the social welfare in the optimum, denoted by W^o , reads

$$W^o = \begin{cases} \frac{2\gamma}{1+\gamma} \frac{(\alpha-c)^{1+\frac{1}{\gamma}} n}{\beta^{\frac{1}{\gamma}}} & \text{if } \frac{(\alpha-c)n^\gamma}{\beta} < \frac{1}{2^\gamma} \\ \alpha - c - \frac{\beta}{2^\gamma(1+\gamma)n^\gamma} & \text{if } \frac{(\alpha-c)n^\gamma}{\beta} \geq \frac{1}{2^\gamma} \end{cases}$$

where the first (resp. second) line corresponds to the case where not all the market (resp. the whole market) is covered in the optimum and $\gamma = 1, 2$. The percentage of welfare losses, defined as before, equals

$$PWL = \begin{cases} 1 - \frac{2+\gamma}{(1+\gamma)^{1+\frac{1}{\gamma}}} & \text{if } \frac{(\alpha-c)n^\gamma}{\beta} < \frac{1}{2^\gamma} \\ 1 - \frac{\frac{2\gamma(2+\gamma)}{(1+\gamma)^{2+\frac{1}{\gamma}}} \left(\frac{(\alpha-c)n^\gamma}{\beta} \right)^{\frac{1}{\gamma}}}{1 - \frac{1}{2^\gamma(1+\gamma)} \frac{\beta}{(\alpha-c)n^\gamma}} & \text{if } \frac{1}{2^\gamma} \leq \frac{(\alpha-c)n^\gamma}{\beta} < \frac{1+\gamma}{2^\gamma} \\ 0 & \text{if } \frac{(\alpha-c)n^\gamma}{\beta} \geq \frac{1+\gamma}{2^\gamma} \end{cases} \quad (37)$$

where $\gamma = 1, 2$. Notice that the first two lines (37) refer to local monopolistic equilibrium. The first (resp. second) line refers to the case in which the market should not (resp. should) be covered in the optimum. The last line refers to kinked and competitive equilibria. Since in the Salop model firms are located optimally and the price just transfers money from consumers to producers, positive welfare losses are only possible when the market is not

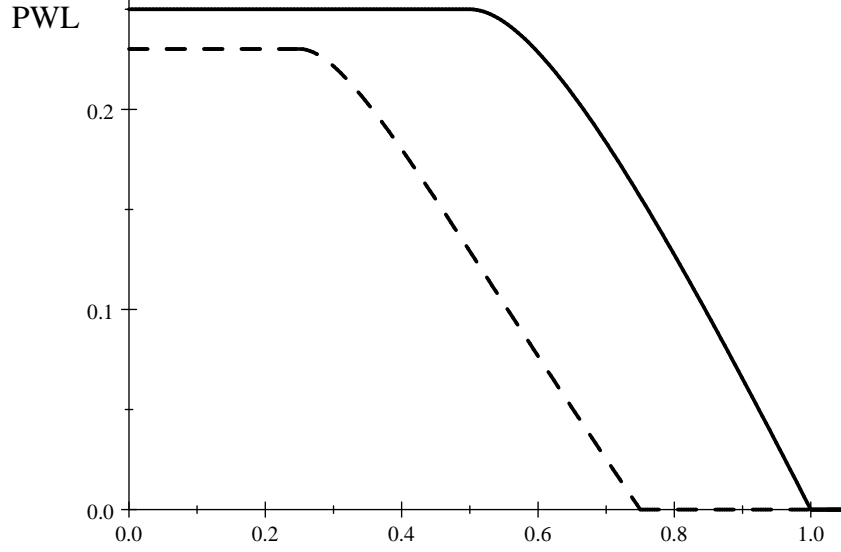


Figure 13: PWL as a function of $\frac{(\alpha-c)n^\gamma}{\beta}$ for linear $\gamma = 1$ (solid) and quadratic $\gamma = 2$ (dash) transportation costs.

covered. Figure 13 depicts PWL as a function of $\frac{(\alpha-c)n^\gamma}{\beta}$ for linear $\gamma = 1$ (solid) and quadratic $\gamma = 2$ (dash) transportation costs.

We now study the relationship between the observable variables and PWL . As before, we assume that market price, outputs, marginal costs, number of active firms and γ can be observed (in the Hotelling model we assumed that $\gamma = 2$). Our view is that γ reflects, basically, the technology of transportation and that this technology is common knowledge.

Formally, let $\{\mathfrak{p}, \mathfrak{x}, \mathfrak{c}, \mathfrak{n}, \gamma\}$ be an *Observation*, where \mathfrak{p} ($> \mathfrak{c}$) stands for market price, \mathfrak{x} is quantity sold by each firm, which is defined as a proportion of consumers purchasing from each firm, \mathfrak{c} is marginal cost, $\mathfrak{n} \in \mathbb{N}$ is number of active firms and $\gamma \in \{1, 2\}$ is the transportation costs. Then, we have the following result.

Proposition 31 *Given an observation $\{\mathfrak{p}, \mathfrak{x}, \mathfrak{c}, \mathfrak{n}, \gamma\}$ there is a Circular Market $\{\alpha, \beta, \mathfrak{c}, \mathfrak{n}, \gamma\}$ such that $\{\mathfrak{p}, \mathfrak{x}\}$ is a Local Monopolistic Equilibrium for this market when not all the market*

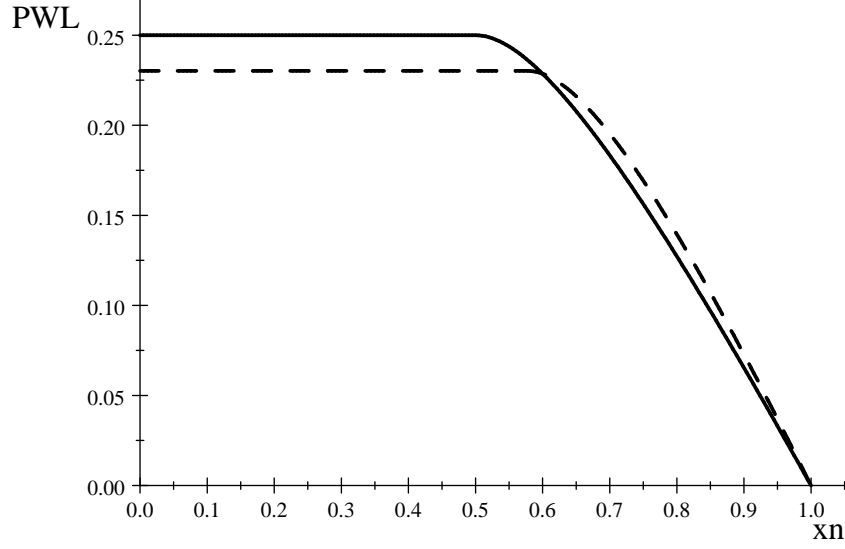


Figure 14: PWL as a function of market coverage $\mathfrak{r}\mathfrak{n}$ for linear (solid) and quadratic (dash) transportation costs.

is covered, i.e. when $\mathfrak{r}\mathfrak{n} < 1$, and PWL is given by

$$PWL = \begin{cases} 1 - \frac{2+\gamma}{(1+\gamma)^{1+\frac{1}{\gamma}}} & \text{if } \mathfrak{r}\mathfrak{n} < \frac{1}{(1+\gamma)^{\frac{1}{\gamma}}} \\ 1 - \frac{\gamma(2+\gamma)\mathfrak{r}\mathfrak{n}}{(1+\gamma)^2 - \frac{1}{(\mathfrak{r}\mathfrak{n})^\gamma}} & \text{if } \frac{1}{(1+\gamma)^{\frac{1}{\gamma}}} \leq \mathfrak{r}\mathfrak{n} < 1 \\ 0 & \text{if } \mathfrak{r}\mathfrak{n} = 1 \end{cases} \quad (38)$$

Figure 14 depicts PWL as a function of observables $\mathfrak{r}\mathfrak{n}$ for linear (solid) and quadratic (dash) transportation costs. As in the Hotelling model, PWL decreases with market coverage, unless in the optimum the whole market should not be covered; in this case PWL is constant (25% under linear transportation costs, as in the standard monopoly model with linear demand). When the whole market is covered both in equilibrium and in the optimum, there are no welfare losses since, as we mentioned above, social welfare does depend on the market coverage and not on prices. Other points are worth discussion.

Firstly, PWL can be calculated in all cases. The indeterminacy that occurred in the Hotelling model does not arise here.

Secondly, as in the Hotelling model and for the same reason, PWL is independent of demand elasticities and markups.

Thirdly, since here there are no misallocations due to wrong locations, PWL is smaller than in the Hotelling model. But these losses may be large.

Finally, Figure 14 shows that PWL is not monotonic in γ for given market coverage. The reason is that a change in γ changes welfare both in equilibrium and in the optimal allocations.

Entry can be considered in this framework by assuming that each firm incurs a fixed cost of entry, K .¹³ In this case PWL is a (non-monotonic) function of $\frac{\alpha-c}{\sqrt{\beta K}}$. PWL can be very large, up to 50%. Given an observation there is a Circular Market with free entry such that the observation is an equilibrium for this market, and $PWL \in (0, \frac{1}{2}]$. Thus, the introduction of fixed costs makes it impossible to infer welfare losses from observations. As before, knowledge of the demand elasticity or the cross elasticity of demand adds nothing. All these results agree with those obtained in Corchón (2008) for the case of Cournot equilibrium with free entry and product homogeneity.

4.4 Vertical Differentiation

In this section we study oligopolistic competition under quality differentiation. This model was developed by Gabszewicz and Thisse (1979, 1980) and Shaked and Sutton (1982, 1983). We consider the simplified version of Shaked and Sutton (1982). Again we have a two-stage game in which in the first stage firms compete in quality (one per firm) and in the second stage they compete in prices.

Consumers' preferences are described by $U = ts - p$ if the consumer purchases one unit of quality s at price p , and by 0 otherwise. The parameter t of taste for quality is uniformly distributed across the population of consumers, $t \sim U[a, b]$ with $0 < a < b$. The density is $\frac{1}{b-a}$.

Assume that there are two firms in the market. Firm $i = 1, 2$ produces a good of quality s_i , where without loss of generality $s_2 > s_1$. Suppose further that s_i must belong to $(0, S]$. We assume zero costs. In particular, the choice of quality is costless.

Definition 32 *A Linear Vertical Market is a list $\{a, b, S\}$ with $b > a > 0$ and $S > 0$.*

¹³The results reported in this paragraph are available upon request.

Consider price competition. Consumers with high taste for quality buy the high-quality good and consumers with lower taste for quality buy the low-quality good (which must be priced lower to attract any consumer), while consumers with the lowest taste for quality might not buy at all. A consumer with taste parameter t_1 is indifferent between purchasing from firm 1 and not purchasing the differentiated commodity if and only if $t_1 s_1 - p_1 = 0$, so $t_1 = \frac{p_1}{s_1}$. A consumer with taste parameter t_2 is indifferent between the two brands if and only if $t_2 s_1 - p_1 = t_2 s_2 - p_2$, so $t_2 = \frac{p_2 - p_1}{s_2 - s_1}$. Therefore, the demand functions read

$$\begin{aligned} D_1 &= \frac{1}{b-a} (t_2 - \max\{a, t_1\}) = \frac{1}{b-a} \left(\frac{p_2 - p_1}{s_2 - s_1} - \max\left\{a, \frac{p_1}{s_1}\right\} \right) \\ D_2 &= \frac{1}{b-a} (b - t_2) = \frac{1}{b-a} \left(b - \frac{p_2 - p_1}{s_2 - s_1} \right) \end{aligned}$$

Each firm i maximizes its profit $\pi_i = p_i D_i(p_i, p_j)$ with respect to p_i . We consider two possible cases in turn: where the market is not covered (i.e. $t_1 > a$) and where the market is covered (i.e. $t_1 \leq a$).

Case where the market is not covered. When $t_1 > a$, some consumers with low taste for quality purchase neither good. Firms' profit maximization yields

$$p_1 = \frac{bs_1(s_2 - s_1)}{4s_2 - s_1} \text{ and } p_2 = \frac{2bs_2(s_2 - s_1)}{4s_2 - s_1}$$

Then profits read

$$\pi_1 = \frac{b^2 s_1 s_2 (s_2 - s_1)}{(b-a)(4s_2 - s_1)^2} \text{ and } \pi_2 = \frac{4b^2 s_2^2 (s_2 - s_1)}{(b-a)(4s_2 - s_1)^2}$$

The condition $t_1 > a$ amounts to $0 < s_1 < \frac{b-4a}{b-a} s_2$ for $b \geq 4a$.

Case where the market is covered. When $t_1 \leq a$, the market is covered and the consumer with lowest taste parameter weakly prefers to purchase product 1. Firms' profit maximization yields

$$p_1 = \frac{1}{3} (s_2 - s_1) (b - 2a) \text{ and } p_2 = \frac{1}{3} (s_2 - s_1) (2b - a),$$

where $b > 2a$ (for both firms compete for consumers). Then profits read

$$\pi_1 = \frac{(b-2a)^2 (s_2 - s_1)}{9(b-a)} \text{ and } \pi_2 = \frac{(2b-a)^2 (s_2 - s_1)}{9(b-a)}$$

Thus, the high-quality firm charges a higher price than the low-quality producer. It also makes a higher profit. For the whole market to be covered in equilibrium, the consumer with taste parameter a should weakly prefer a low-quality good to nothing, i.e. $as_1 - p_1 \geq 0$, which amounts to $\frac{b-2a}{a+b}s_2 \leq s_1 < s_2$ for $b > 2a$.

Summing up the results of price competition, if $2a < b < 4a$ an equilibrium arises where the entire market is covered. If $b \geq 4a$ the market might be not covered and two types of equilibria arise: an equilibrium with uncovered market for $0 < s_1 < \frac{b-4a}{b-a}s_2$ and an equilibrium with covered market for $\frac{b-2a}{a+b}s_2 \leq s_1 < s_2$.

In the first stage, each firm i maximizes $\pi_i(s_i, s_j)$ over s_i . The "marginal relocation tendency" of firm 2 reads $\frac{\partial \pi_2}{\partial s_2} \geq 0$ for all s_1 and $b > 2a$, therefore firm 2 chooses the maximal quality level $s_2^* = S$. Firm 1's profit maximization yields

$$s_1^* = \begin{cases} \frac{b-2a}{a+b}S & \text{if } 2a < b \leq 8a \\ \frac{4}{7}S & \text{if } b > 8a. \end{cases}$$

Figure 15 depicts $\frac{s_1^*}{S}$ as a function of $\frac{b}{a}$. Note that for $2 < \frac{b}{a} \leq 8$ the entire market is covered, and the consumer with the lowest taste parameter a is indifferent between buying the low quality product and neither product. The larger the market, the higher the quality of good 1. The reason is that price competition between two goods drives their prices down to a level at which not even the consumer with the lowest taste for quality would want to buy good 1 if its quality is very low. So to attract consumers, firm 1 has to raise the quality of its good. At some point however—at $\frac{b}{a} = 8$ exactly—firm 1 prefers to freeze the quality at constant level $s_1^* = \frac{4}{7}S$ and not to serve consumers with low taste parameters such that the market becomes uncovered. Thus, the discontinuity arises at the point in which the market becomes uncovered.

Definition 33 *An equilibrium in the linear vertical market $\{a, b, S\}$ with $b > 2a$ is a list of qualities $\{s_1^*, s_2^*\}$, a list of prices $\{p_1^*, p_2^*\}$ and the percentage of market coverage m^* such that*

$$\{s_1^*, s_2^*, p_1^*, p_2^*, m^*\} = \begin{cases} \left\{ \frac{b-2a}{a+b}S, S, \frac{a(b-2a)}{a+b}S, \frac{a(2b-a)}{a+b}S, 1 \right\} & \text{if } 2a < b \leq 8a \\ \left\{ \frac{4}{7}S, S, \frac{1}{14}Sb, \frac{1}{4}Sb, \frac{7b}{8(b-a)} \right\} & \text{if } b > 8a. \end{cases}$$

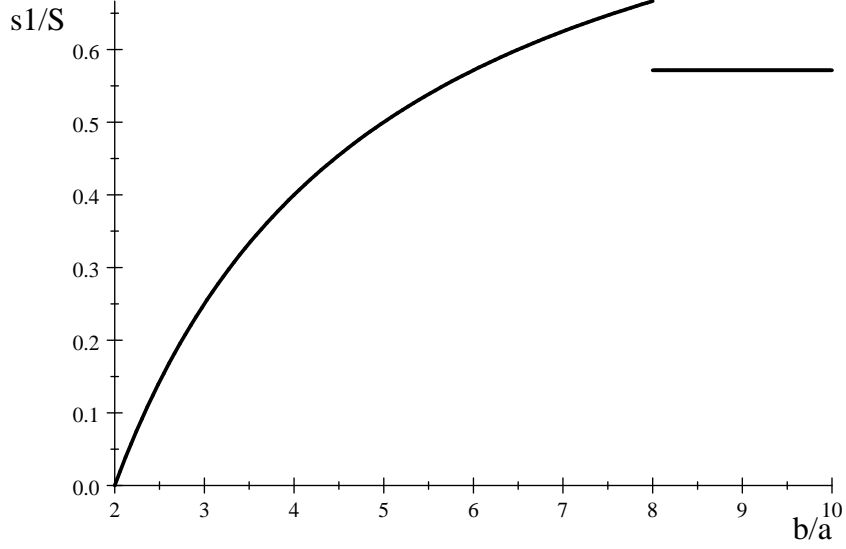


Figure 15: $\frac{s_1^*}{S}$ as a function of $\frac{b}{a} > 2$.

Define social welfare, denoted by W , as the gross consumers' surplus:

$$W = \int_{\max\{a, t_1\}}^{t_2} \frac{s_1 t}{b-a} dt + \int_{t_2}^b \frac{s_2 t}{b-a} dt,$$

which in the equilibrium reads

$$W^* = \begin{cases} \frac{(5a^3 - 5a^2b + 2ab^2 + 3b^3)}{6(b^2 - a^2)} S & \text{if } 2a < b \leq 8a \\ \frac{11b^2}{24(b-a)} S & \text{if } b > 8a. \end{cases}$$

A social planner would choose the brands' quality to maximize the social welfare, thus $s_1^o = s_2^o = S$. Hence in the optimal allocation there are two undifferentiated firms that make no profit, and the social welfare reads

$$W^o = \frac{(a+b)}{2} S.$$

Consequently, the percentage of welfare losses reads

$$PWL = 1 - \frac{W^*}{W^o} = \begin{cases} \frac{(\frac{b}{a}-2)(4+\frac{b}{a})}{3(\frac{b}{a}-1)(1+\frac{b}{a})^2} & \text{if } 2 < \frac{b}{a} \leq 8 \\ \frac{(\frac{b}{a})^2 - 12}{12(\frac{b}{a}-1)(\frac{b}{a}+1)} & \text{if } \frac{b}{a} > 8. \end{cases} \quad (39)$$

Figure 16 depicts PWL as a function of $\frac{b}{a}$. Notice that welfare losses are not large, with a maximum of about 8.33% (which is reached for very large markets, i.e., for $\frac{b}{a} \rightarrow \infty$).

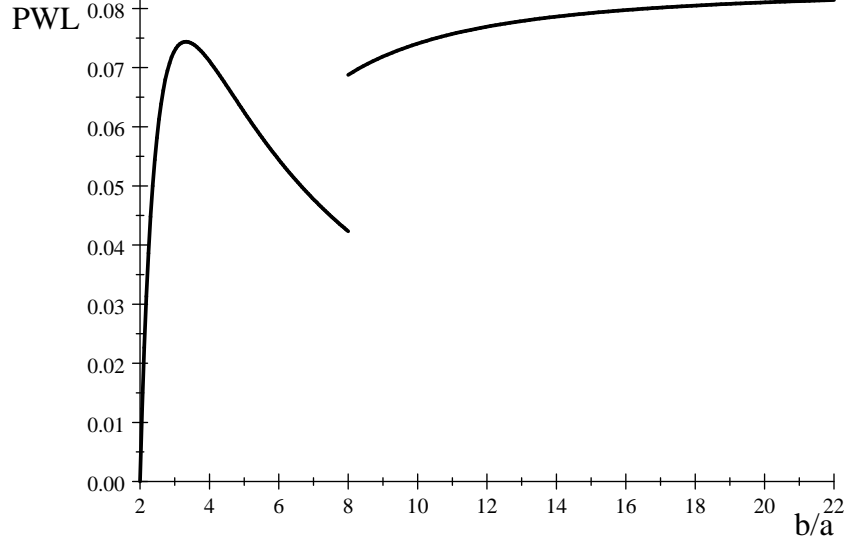


Figure 16: PWL as a function of $\frac{b}{a} > 2$.

We are interested in PWL yielded by this market, conditional on the values taken by certain variables that can be observed, namely market prices and market coverage. Formally:

Definition 34 An observation is a list $\{p_1, p_2, m\}$ where $p_1 > 0$ is low-quality good price, $p_2 > p_1$ is high-quality good price and $m \in [0, 1]$ is the percentage of market coverage.

Let us relate PWL with observable variables.

Proposition 35 i) Given an observation $\{p_1, p_2, m\}$ where $m = 1$ and $0 < \frac{p_1}{p_2} \leq \frac{2}{5}$, there is a linear vertical market $\{a, b, S\}$ such that $\left\{S_{\frac{p_1}{p_2 - p_1}}, S, p_1, p_2, m\right\}$ is an equilibrium for this market, and

$$PWL = \frac{\frac{p_1}{p_2} \left(6 \left(\frac{p_1}{p_2} \right)^2 - 7 \frac{p_1}{p_2} + 2 \right)}{3 \left(1 - \frac{p_1}{p_2} \right)^2 \left(\frac{p_1}{p_2} + 1 \right)} \quad (40)$$

ii) Given an observation $\{p_1, p_2, m\}$ where $\frac{7}{8} < m < 1$, and $\frac{p_1}{p_2} = \frac{2}{7}$, there is a linear vertical market $\{a, b, S\}$ such that $\left\{\frac{4}{7}S, S, p_1, p_2, m\right\}$ is an equilibrium for this market, and

$$PWL = \frac{\left(\frac{8m}{8m-7} \right)^2 - 12}{12 \left(\frac{8m}{8m-7} - 1 \right) \left(\frac{8m}{8m-7} + 1 \right)} \quad (41)$$

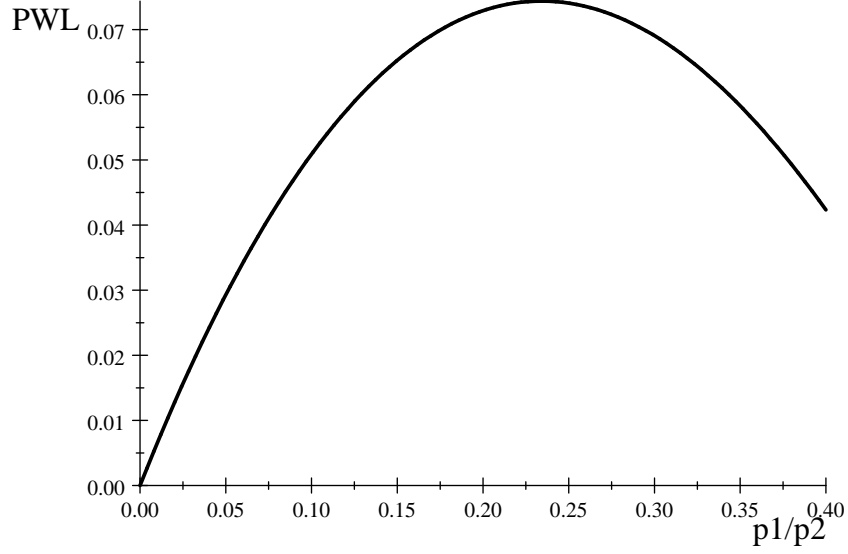


Figure 17: PWL as a function of $\frac{p_1}{p_2}$ for $m = 1$ and $0 < \frac{p_1}{p_2} \leq \frac{2}{5}$.

Figure 17 depicts PWL as a function of $\frac{p_1}{p_2}$ for $m = 1$ and $0 < \frac{p_1}{p_2} \leq \frac{2}{5}$. PWL is easily seen to be single-peaked in the domain $(0, \frac{2}{5}]$ reaching a maximum at, approximately, $\frac{p_1}{p_2} = 0.23$. Thus, an increase in relative prices can decrease or increase relative welfare losses. Figure 18 depicts PWL as a function of m for $\frac{7}{8} < m < 1$ and $\frac{p_1}{p_2} = \frac{2}{7}$. Here PWL depends only on the market coverage, in a decreasing way, as expected.

4.5 Conclusion

In this paper we have shown that the task of finding welfare losses from market data in models of horizontal (i.e. location) and vertical (i.e. quality) differentiation is feasible, provided that a specific functional form is assumed.

In location models welfare losses can be very high. Welfare losses are due mainly to lack of market cover and/or firms located in the wrong positions. To the best of our knowledge, this point has never been recognized by the literature on empirical measurement of welfare losses. Markups and demand elasticities do not play any role in determining welfare losses in the location models considered here. They might play a role if consumers were allowed to buy several goods.

Under vertical differentiation welfare losses can be read from prices and market coverage

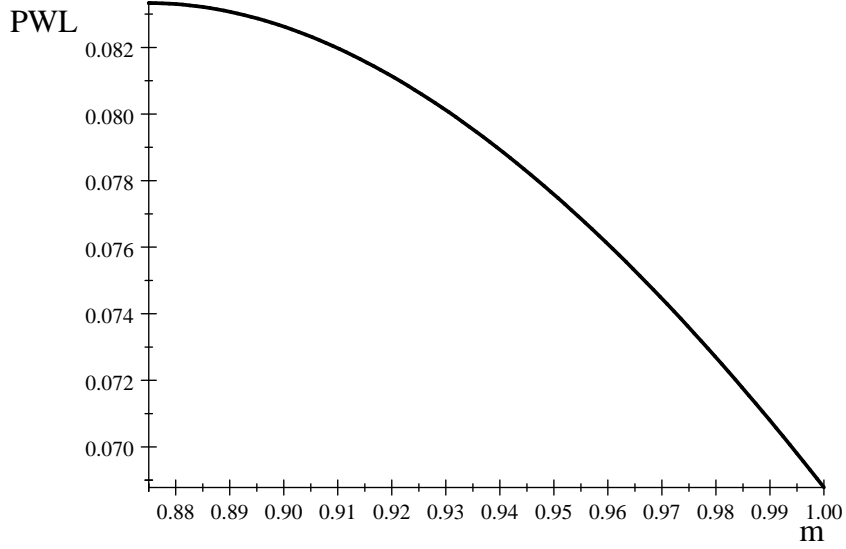


Figure 18: PWL as a function of m for $\frac{7}{8} < m < 1$ and $\frac{p_1}{p_2} = \frac{2}{7}$.

but they are discontinuous. The percentage of welfare losses here is not very large. Thus, despite the apparent similarities in the derivation of equilibrium in models of horizontal and vertical differentiation, these two strands of models are very different from the point of view of welfare losses.

The models considered in this paper are symmetric and rely on specific forms of the commodity space. We hope that our findings can be used to tackle models with asymmetric firms, see e.g. Aghion and Schankerman (2004) for a Salop-like model with heterogeneous costs or with other forms of the commodity space such as the Spokes model of Chen and Riordan (2007). Another possible extension of our work would be to study consumer and producer surpluses separately.

4.6 Appendix

Proof of Proposition 29: First, consider the case where not all the market is covered, i.e. $m < 1$ and $\mathfrak{r} < \frac{1}{4}$. Let $\alpha = \frac{5p-c}{4}$ and $\beta = \frac{p-c}{4\mathfrak{r}^2}$. We easily see that the linear horizontal market $\{\alpha, \beta, c\}$ yields a local monopolistic equilibrium where $p_m^* = p$ and $\{x_{m1}^*, x_{m2}^*\} = \{\mathfrak{r}, 1 - \mathfrak{r}\}$. When $\mathfrak{r} < \frac{1}{4\sqrt{5}}$ in the optimum not all the market is covered and $PWL = 1 - \frac{7}{5\sqrt{5}}$. When $\mathfrak{r} \geq \frac{1}{4\sqrt{5}}$ in the optimum the entire market should be covered. Plugging $\alpha = \frac{5p-c}{4}$ and

$\beta = \frac{\mathfrak{p}-\mathfrak{c}}{4\mathfrak{r}^2}$ in the second line of (35) yields the second line of (36) for PWL as a function of observables.

Next, consider the case where the entire market is covered, i.e. $\mathfrak{m} = 1$, and the firms do not locate at the edges of the market, i.e. $0 < \mathfrak{r} \leq \frac{1}{4}$. Let $\alpha = \frac{1}{2}(3\mathfrak{p} - \mathfrak{c} - 2\mathfrak{r}(\mathfrak{p} - \mathfrak{c}))$ and $\beta = \frac{2(\mathfrak{p}-\mathfrak{c})}{1-2\mathfrak{r}}$. It is straightforward to check that the linear horizontal market $\{\alpha, \beta, \mathfrak{c}\}$ yields a kinked equilibrium where $p_k^* = \mathfrak{p}$ and $\{x_{k1}^*, x_{k2}^*\} = \{\mathfrak{r}, 1 - \mathfrak{r}\}$. Plugging $\alpha = \frac{1}{2}(3\mathfrak{p} - \mathfrak{c} - 2\mathfrak{r}(\mathfrak{p} - \mathfrak{c}))$ and $\beta = \frac{2(\mathfrak{p}-\mathfrak{c})}{1-2\mathfrak{r}}$ in the third line of (35) yields the third line of (36) for PWL as a function of observables.

Finally, let us consider the case where the entire market is covered, i.e. $\mathfrak{m} = 1$, and the firms locate at the edges of the market, i.e. $\mathfrak{r} = 0$. Here, with available observables there is no way to distinguish between the kinked equilibrium where the firms are located at the market edges and the competitive equilibrium.

In the case of the kinked equilibrium where the firms are located at the market edges, let us fix $\frac{5\mathfrak{p}-\mathfrak{c}}{4} < \alpha \leq \frac{3\mathfrak{p}-\mathfrak{c}}{2}$ (for condition $\frac{3}{4} \leq \frac{\alpha-\mathfrak{c}}{\beta} < \frac{5}{4}$ to hold) and let $\beta = 4(\alpha - \mathfrak{p})$. It is straightforward to check that the linear horizontal market $\{\alpha, \beta, \mathfrak{c}\}$ yields a kinked equilibrium where $p_k^* = \mathfrak{p}$ and $\{x_{k1}^*, x_{k2}^*\} = \{0, 1\}$. From the fourth line of (35) we get PWL in the kinked equilibrium, denoted as PWL_k , as a function of observables and α :

$$PWL_k = \frac{3(\alpha - \mathfrak{p})}{11\alpha + \mathfrak{p} - 12\mathfrak{c}} \text{ where } \frac{5\mathfrak{p} - \mathfrak{c}}{4} < \alpha \leq \frac{3\mathfrak{p} - \mathfrak{c}}{2}$$

which is increasing in α and achieves its maximal value of $\frac{3}{35}$ at $\alpha = \frac{3\mathfrak{p}-\mathfrak{c}}{2}$ and its minimal value of $\frac{3}{59}$ at $\alpha = \frac{5\mathfrak{p}-\mathfrak{c}}{4}$. This and the continuity of PWL_k with respect to α imply that $PWL_k \in (\frac{3}{59}, \frac{3}{35}]$.

In the case of the competitive equilibrium, let us fix $\alpha \geq \frac{5\mathfrak{p}-\mathfrak{c}}{4}$ (for condition $\frac{\alpha-\mathfrak{c}}{\beta} \geq \frac{5}{4}$ to hold) and let $\beta = \mathfrak{p} - \mathfrak{c}$. It is easy to show that the linear horizontal market $\{\alpha, \beta, \mathfrak{c}\}$ yields a competitive equilibrium where $p_c^* = \mathfrak{p}$ and $\{x_{c1}^*, x_{c2}^*\} = \{0, 1\}$. From the fourth line of (35) we get PWL in the competitive equilibrium, denoted as PWL_c , as a function of observables and α :

$$PWL_c = \frac{3(\mathfrak{p} - \mathfrak{c})}{48\alpha - 47\mathfrak{c} - \mathfrak{p}} \text{ where } \alpha \geq \frac{5\mathfrak{p} - \mathfrak{c}}{4}$$

which is decreasing in α and achieves its maximal value of $\frac{3}{59}$ at $\alpha = \frac{5\mathfrak{p}-\mathfrak{c}}{4}$ and goes to

0 as α goes to infinity. This and the continuity of PWL_c with respect to α imply that $PWL_c \in (0, \frac{3}{59}]$. Thus, in the case where the entire market is covered and the firms locate at the edges of the market, i.e. $\mathfrak{m} = 1$ and $\mathfrak{r} = 0$, $PWL \in (0, \frac{3}{35}]$ with $\frac{3}{35} \approx 0.0857$. ■

Proof of Proposition 31: When $\mathfrak{r} < \frac{1}{\mathfrak{n}}$ (that is, not all the market is covered), let $\alpha = \frac{(1+\gamma)\mathfrak{p}-\mathfrak{c}}{\gamma}$ and $\beta = \frac{2\gamma}{\gamma} \frac{\mathfrak{p}-\mathfrak{c}}{\mathfrak{r}^\gamma}$. We easily see that the circular market $\{\alpha, \beta, \mathfrak{c}, \mathfrak{n}, \gamma\}$ yields a local monopolistic equilibrium where $p_m^* = \mathfrak{p}$ and $D_i(p_m^*) = \mathfrak{r}$. When $\mathfrak{r}\mathfrak{n} \geq \frac{1}{(1+\gamma)^{\frac{1}{\gamma}}}$ in the optimum the whole market should be covered. Plugging $\alpha = \frac{(1+\gamma)\mathfrak{p}-\mathfrak{c}}{\gamma}$ and $\beta = \frac{2\gamma}{\gamma} \frac{\mathfrak{p}-\mathfrak{c}}{\mathfrak{r}^\gamma}$ in (37) yields the formula (38) for PWL as a function of observable market coverage $\mathfrak{r}\mathfrak{n}$ and transportation costs γ . ■

Proof of Proposition 35: First, consider the case where the entire market is covered, $\mathfrak{m} = 1$, and $0 < \frac{\mathfrak{p}_1}{\mathfrak{p}_2} \leq \frac{2}{5}$. Let us fix $S > 0$ and let

$$\mathfrak{p}_1 = \frac{a(b-2a)}{a+b}S, \mathfrak{p}_2 = \frac{a(2b-a)}{a+b}S,$$

which yields

$$a = \frac{\mathfrak{p}_2 - \mathfrak{p}_1}{S}, b = \frac{(2\mathfrak{p}_2 - \mathfrak{p}_1)(\mathfrak{p}_2 - \mathfrak{p}_1)}{(\mathfrak{p}_2 - 2\mathfrak{p}_1)S}$$

It is straightforward to check that $2 < \frac{b}{a} \leq 8$ since $0 < \frac{\mathfrak{p}_1}{\mathfrak{p}_2} \leq \frac{2}{5}$. Then by construction the market $\{a, b, S\}$ yields an equilibrium where $\{s_1^*, s_2^*, p_1^*, p_2^*, m^*\} = \{S \frac{\mathfrak{p}_1}{\mathfrak{p}_2 - \mathfrak{p}_1}, S, \mathfrak{p}_1, \mathfrak{p}_2, \mathfrak{m}\}$. Plugging $a = \frac{\mathfrak{p}_2 - \mathfrak{p}_1}{S}$ and $b = \frac{(2\mathfrak{p}_2 - \mathfrak{p}_1)(\mathfrak{p}_2 - \mathfrak{p}_1)}{(\mathfrak{p}_2 - 2\mathfrak{p}_1)S}$ into the first line of (39) we get (40) for PWL as a function of an observation $\{\mathfrak{p}_1, \mathfrak{p}_2, \mathfrak{m}\}$.

Second, consider the case where the market is not covered, $\frac{7}{8} < \mathfrak{m} < 1$, and $\frac{\mathfrak{p}_1}{\mathfrak{p}_2} = \frac{2}{7}$. Fix $S > 0$, $a > 0$ and let

$$\mathfrak{p}_1 = \frac{1}{14}Sb, \mathfrak{p}_2 = \frac{1}{4}Sb \text{ and } \mathfrak{m} = \frac{7b}{8(b-a)},$$

which yields

$$b = \frac{8\mathfrak{m}}{8\mathfrak{m}-7}a.$$

One can easily check that $\frac{b}{a} > 8$ for $\frac{7}{8} < \mathfrak{m} < 1$. Then by construction the market $\{a, b, S\}$ yields an equilibrium where $\{s_1^*, s_2^*, p_1^*, p_2^*, m^*\} = \{\frac{4}{7}S, S, \mathfrak{p}_1, \mathfrak{p}_2, \mathfrak{m}\}$. Plugging $\frac{b}{a} = \frac{8\mathfrak{m}}{8\mathfrak{m}-7}$ into the second line of (39) yields (41) for PWL as a function of observables. ■

REFERENCES

- [1] Acemoglu, D., Bimpikisb, K. and Ozdaglar, A. (2009). "Price and Capacity Competition." *Games and Economic Behavior*, 66, 1, 1-26.
- [2] Aghion, P. and Schankerman, M. (2004). "On the Welfare Effects and Political Economy of Competition-Enhancing Policies." *The Economic Journal*, 114, 800-824.
- [3] Alesina, A. and Rosenthal, H. (1995). *Partisan Politics, Divided Government, and the Economy*. Cambridge: Cambridge University Press.
- [4] Alesina, A. and Rosenthal, H. (1996). "A Theory of Divided Government." *Econometrica*, 64, 1311-1341.
- [5] Alesina, A. and Tabellini, G. (2008). "Bureaucrats or Politicians? Part II: Multiple Policy Tasks." *Journal of Public Economics*, 92, 426-447.
- [6] Amir, R. and Jin, J.Y. (2001). "Cournot and Bertrand Equilibria Compared: Substitutability, Complementarity and Concavity." *International Journal of Industrial Organization*, 19, 303-317.
- [7] Anderson, S.P. and Renault, R. (2003). "Efficiency and Surplus Bounds in Cournot Competition." *Journal of Economic Theory*, 113, 253-264.
- [8] Austen-Smith, D. (1997). "Interest Groups: Money, Information, and Influence." In *Perspectives on Public Choice. A Handbook*, edited by D. C. Mueller. Cambridge University Press.
- [9] Austen-Smith, D. and Banks, J.S. (1989). "Electoral Accountability and Incumbency." In *Models of Strategic Choice in Politics*, edited by P. C. Ordeshook. Ann Arbor MI: University of Michigan Press.
- [10] Banks, J.S. and Sundaram, R.K. (1993). "Adverse Selection and Moral Hazard in a Repeated Elections Model," in *Political Economy: Institutions, Information, Competition and Representation*, edited by W. Barnett, M. Hinich, and N. Schofield. New York: Cambridge University Press.
- [11] Banks, J.S. and Sundaram, R.K. (1996). "Electoral Accountability and Selection Effects." University of Rochester, mimeo.
- [12] Baron, D. (1989). "Service-Induced Campaign Contributions and the Electoral Equilibrium." *Quarterly Journal of Economics*, 104, 45-72.
- [13] Barro, R. (1973). "The Control of Politicians: An Economic Model." *Public Choice*, 14, 19-42.
- [14] Becker, G.S. (1983). "A Theory of Competition among Pressure Groups for Political Influence." *Quarterly Journal of Economics*, 98, 371-400.

- [15] Becker, G.S. (1985). "Public Policies, Pressure Groups and Deadweight Costs." *Journal of Public Economics*, 28, 330-347.
- [16] Becker, G., DeGroot, M. and Marschak, J. (1964). "Measuring Utility by a Single-response Sequential Method." *Behavioral Science*, 9, 226-236.
- [17] Belleflamme, P. and Hindriks, J. (2005). "Yardstick Competition and Political Agency Problems." *Social Choice and Welfare*, 24, 155-169.
- [18] Bernheim, B.D. and Whinston, M.D. (1986). "Menu Auctions, Resource Allocation, and Economic Influence." *Quarterly Journal of Economics*, 101, 1-31.
- [19] Besley, T. (2006). *Principled Agents? The Political Economy of Good Government*. Oxford University Press.
- [20] Besley, T. and Case, A. (1995a). "Incumbent Behavior: Vote Seeking, Tax Setting and Yardstick Competition." *American Economic Review*, 85, 25-45.
- [21] Besley, T. and Case, A. (1995b). "Does Political Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits." *Quarterly Journal of Economics*, 110, 769-798.
- [22] Besley, T. and Case, A. (2003). "Political Institutions and Policy Choices: Evidence from the United States." *Journal of Economic Literature*, 41, 7-73.
- [23] Besley, T. and Coate, S. (1997). "An Economic Model of Representative Democracy." *Quarterly Journal of Economics*, 112, 85-114.
- [24] Besley, T. and Smart, M. (2007). "Fiscal Restraints and Voter Welfare." *Journal of Public Economics*, 91, 755-773.
- [25] Bolton, P. and Dewatripont, M. (2005). *Contract Theory*. MIT Press.
- [26] Bombardini, M. (2008). "Firm Heterogeneity and Lobby Participation." *Journal of International Economics*, 75, 329-348.
- [27] Bordignon, M., Cerniglia, F. and Revelli, F. (2004). "Yardstick Competition in Intergovernmental Relationships: Theory and Empirical Predictions." *Economics Letters*, 83, 325-333.
- [28] Breton, A. (1996). *Competitive Governments: an Economic Theory of Politics and Public Finance*. Cambridge and New York: Cambridge University Press.
- [29] Breton, A. and Frascini, A. (2003). "Vertical Competition in Unitary States: the Case of Italy." *Public Choice*, 114, 57-77.
- [30] Breton, A. and Salmon, P. (2001). "External Effects of Domestic Regulations: Comparing Internal and International Barriers to Trade." *International Review of Law and Economics*, 21, 135-155.
- [31] Bugarin, M.S. (2003). "Vote Splitting, Reelection and Electoral Control: Towards a Unified Model." *Social Choice and Welfare*, 20, 137-154.

- [32] Burden, B.C. and Kimball, D.C. (2002). *Why Americans Split Their Tickets: Campaigns, Competition, and Divided Government*. University of Michigan Press, Ann Arbor.
- [33] Cable, J., Carruth, A. and Dixit, A. (1994). "Oligopoly and Welfare." In *Current Issues in Industrial Economics*, edited by J. Cable. Macmillan Press: London.
- [34] Campante, F.R. and Ferreira, F.H.G. (2007). "Inefficient lobbying, populism and oligarchy." *Journal of Public Economics*, 91, 993–1021.
- [35] Chari, V. V., Jones, L.E. and Marimon, R. (1997). "The Economics of Split-Ticket Voting in Representative Democracies." *American Economic Review*, 87, 957-976.
- [36] Chen, Y. and Riordan, M. H. (2007). "Price and Variety in the Spokes Model." *The Economic Journal*, 117, 897-921.
- [37] Coate, M. (2005). "Empirical Analysis of Merger Enforcement Under the 1992 Merger Guidelines." *Review of Industrial Organization*, 27, 4, 279-301.
- [38] Corchón, L. C. (2008). "Welfare Losses under Cournot Competition." *International Journal of Industrial Organization*, 26, 1120-1131.
- [39] Corchón, L. C. and Zudenkova, G. (2008). "Computing Welfare Losses from Data under Imperfect Competition with Heterogeneous Goods." UC3M WP 08-26 (16).
- [40] Corchón, L. C. and Zudenkova, G. (2008). "Computing Welfare Losses from Data under Imperfect Competition with Heterogeneous Goods." *International Journal of Industrial Organization*, 27, 646-654.
- [41] Cowling, K. and Mueller, D.C. (1978). "The Social Costs of Monopoly Power." *Economic Journal*, 88, 727-748.
- [42] Damania, R. and Fredriksson, P.G. (2000). "On the Formation of Industry Lobby Groups." *Journal of Economic Behavior & Organization*, 41, 315-335.
- [43] Damania, R. and Fredriksson, P.G. (2003). "Trade Policy Reform, Endogenous Lobby Group Formation, and Environmental Policy." *Journal of Economic Behavior & Organization*, 52, 47-69.
- [44] D'Aspremont, C., Gabszewicz, J. J. and Thisse, J.-F. (1979). "On Hotelling's Stability in Competition." *Econometrica*, 47, 1145-1150.
- [45] Daughety, A. (1990). "Beneficial concentration." *American Economic Review*, 80, 1231-1237.
- [46] Dixit, A. (1979). "A Model of Duopoly Suggesting a Theory of Entry Barriers." *Bell Journal of Economics*, 10, 20-32.
- [47] Dixit, A., Grossman, G.M. and Helpman, E. (1997). "Common Agency and Coordination: General Theory and Application to Government Policy Making." *Journal of Political Economy*, 105, 752-769.
- [48] Dixit, A. and Stiglitz, J. E. (1977). "Monopolistic Competition and Optimum Product Diversity." *American Economic Review*, 67, 297-308.

- [49] Domhoff, G.W. (1983). *Who Rules America now?* Englewood Cliffs, NJ: Prentice-Hall.
- [50] Drazen, A., Limão, N. and Stratmann, T. (2007). "Political Contribution Caps and Lobby Formation: Theory and Evidence." *Journal of Public Economics*, 91, 723–754.
- [51] Economides, N. (1984). "The Principle of Minimum Differentiation Revisited." *European Economic Review*, 24, 345-368.
- [52] Economides, N. (1986). "Minimal and Maximal Product Differentiation in Hotelling's Duopoly." *Economics Letters*, 21, 67-71.
- [53] Economides, N. (1989). "Symmetric Equilibrium Existence and Optimality in Differentiated Product Markets." *Journal of Economic Theory*, 47, 178-194.
- [54] Fair, R.C. (1978). "The Effect of Economic Events on Votes for President." *Review of Economics and Statistics*, LX, 159-173.
- [55] Farrell, J. and Shapiro, C. (1990). "Horizontal Merger: an Equilibrium Analysis." *American Economic Review*, 80, 107-126.
- [56] Felli, L. and Merlo, A. (2006). "Endogenous Lobbying." *Journal of the European Economic Association*, 4, 180-215.
- [57] Felli, L. and Merlo, A. (2007). "If You Cannot Get Your Friends Elected Lobby Your Enemies." *Journal of the European Economic Association*, 5, 624-635.
- [58] Ferejohn, J. (1986). "Incumbent Performance and Electoral Control." *Public Choice*, 50, 5-26.
- [59] Ferreira, F. and Gyourko, J. (2009). "Do Political Parties Matter? Evidence from U.S. Cities." *Quarterly Journal of Economics*, 124, 399-422.
- [60] Fiorina, M.P. (1996). *Divided Government*. 2d ed. Needham Heights, MA: Allyn & Bacon.
- [61] Formby, J. and Layson, S. (1982). "Allocative Inefficiency and Measures of Market Power." *Atlantic Economic Journal*, 10, 4, 67-70.
- [62] Fox, J. and Van Weelden, R. (2009). "Partisanship and the Effectiveness of Oversight." Yale University, mimeo.
- [63] Gabszewicz, J. J. and Thisse, J.-F. (1979). "Price Competition, Quality and Income Disparities." *Journal of Economic Theory*, 20, 340-359.
- [64] Gabszewicz, J. J. and Thisse, J.-F. (1980). "Entry (and Exit) in a Differentiated Industry." *Journal of Economic Theory*, 22, 327-338.
- [65] Glazer, A. and Gradstein, M. (2005). "Elections with Contribution-Maximizing Candidates." *Public Choice*, 122, 467-482.
- [66] Grossman, G.M. and Helpman, E. (1994). "Protection for Sale." *American Economic Review*, 84, 833-850.

- [67] Grossman, G.M. and Helpman, E. (1996). "Electoral Competition and Special Interest Politics." *Review of Economic Studies*, 63, 265-286.
- [68] Grossman, G.M. and Helpman, E. (2001). *Special Interest Politics*. MIT Press.
- [69] Häckner, J. (2000). "A Note on Price and Quantity Competition in Differentiated Oligopolies." *Journal of Economic Theory*, 93, 2, 233-239.
- [70] Harberger, A.C. (1954). "Monopoly and Resource Allocation." *American Economic Review: Papers and Proceedings*, 44, 77-87.
- [71] Helpman, E. and Persson, T. (2001). "Lobbying and Legislative Bargaining." *Advances in Economic Analysis & Policy*, 1, Article 3.
- [72] Horowitz, J. (2006). "The Becker-DeGroot-Marschak Mechanism is not Necessarily Incentive Compatible, even for Non-random Goods." *Economics Letters*, 93, 6-11.
- [73] Hotelling, H. (1929). "Stability in Competition." *Economic Journal*, 39, 41-57.
- [74] Hsu, J. and Wang, X.H. (2005). "On Welfare under Cournot and Bertrand Competition in Differentiated Oligopolies." *Review of Industrial Organization*, 27, 2, 185-191.
- [75] Jacobson, G.C. (1990). *The Electoral Origins of Divided Government: Competition in U.S. House Elections, 1946-1988*. Westview Press, Boulder.
- [76] Johari, R. and Tsitsiklis, J. (2005). "Efficiency loss in Cournot games." Mimeo, Jan. 28.
- [77] Kramer, G.H. (1971). "Short-Term Fluctuations in U.S. Voting Behavior, 1896-1964." *American Political Science Review*, 65, 131-143.
- [78] Krusell, P., Quadrini, V. and Ríos-Rull, J.-V. (1997). "Politico-Economic Equilibrium and Economic Growth." *Journal of Economic Dynamics and Control*, 21, 243-272.
- [79] Lahiri, S. and Ono, Y. (1988). "Helping Minor Firms Reduces Welfare." *Economic Journal*, 98, 1199-1202.
- [80] Laussel, D.G. (2006). "Special Interest Politics and Endogenous Lobby Formation." *Topics in Theoretical Economics*, 6, Article 1.
- [81] Lerner, A. (1934). "The Concept of Monopoly and the Measurement of Monopoly Power." *Review of Economic Studies*, 1, 3, 157-175.
- [82] Magee, C. (2002). "Endogenous Trade Policy and Lobby Formation: an Application to the Free-Rider Problem." *Journal of International Economics*, 57, 449-471.
- [83] McCarty, N., Poole, K. and Rosenthal, H. (2006). *Polarized America*. MIT Press.
- [84] Mills, C.W. (1956). *The Power Elite*. New York: Oxford University Press.
- [85] Mitra, D. (1999). "Endogenous Lobby Formation and Endogenous Protection: A Long-Run Model of Trade Policy Determination." *American Economic Review*, 89, 1116-1134.

- [86] McHardy, J.P. (2000). "Miscalculations of Monopoly and Oligopoly Welfare Losses with Linear Demand." Hull Economic Research Papers, November.
- [87] Nisan, N., Roughgarden, T., Tardos, E. and Vazirani, V. (2007). *Algorithmic Game Theory*, Cambridge University Press.
- [88] Olson, M. (1971). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.
- [89] Osborne, M.J. and Slivinski, A. (1996). "A Model of Political Competition with Citizen-Candidates." *Quarterly Journal of Economics*, 111, 65-96.
- [90] Peltzman, S. (1992). "Voters as Fiscal Conservatives." *Quarterly Journal of Economics*, 107, 327-361.
- [91] Persson, T. (1998). "Economic Policy and Special Interest Politics." *Economic Journal*, 108, 310-327.
- [92] Persson, T., Roland, G. and Tabellini, G. (1997). "Separation of Powers and Political Accountability." *Quarterly Journal of Economics*, 112, 1163-1202.
- [93] Persson, T. and Tabellini, G. (2002). *Political Economics: Explaining Economic Policy*. MIT Press.
- [94] Salmon, P. (1987). "Decentralization as an Incentive Scheme." *Oxford Review of Economic Policy*, 3, 24-43.
- [95] Salop, S. C. (1979). "Monopolistic Competition with Outside Goods." *Bell Journal of Economics*, 10, 141-156.
- [96] Sand-Zantman, W. (2004). "Economic Integration and Political Accountability." *European Economic Review*, 48, 1001-1025.
- [97] Shaked, A. and Sutton, J. (1982). "Relaxing Price Competition through Product Differentiation." *Review of Economic Studies*, 49, 3-13.
- [98] Shaked, A. and Sutton, J. (1983). "Natural Oligopolies." *Econometrica*, 51, 1469-1483.
- [99] Singh, N. and Vives, X. (1984). "Price and Quantity Competition in a Differentiated Oligopoly." *Rand Journal of Economics*, 15, 546-554.
- [100] Smith, M.A. (2000). *American Business and Political Power: Public Opinion, Elections, and Democracy*. Chicago: University of Chicago Press.
- [101] Snyder, J. (1990). "Campaign Contributions as Investments: The U.S. House of Representatives, 1980-86." *Journal of Political Economy*, 98, 1195-1227.
- [102] Spence, M. (1976). "Product Selection, Fixed Costs, and Monopolistic Competition." *Review of Economic Studies*, 43, 217-235.
- [103] Stigler, G. J. (1949). "Monopolistic Competition." In *Retrospect in Five Lectures on Economic Problems*. London: Longmans, Green and Co.
- [104] Sutton, J. (1998). *Technology and Market Structure*. Cambridge, MA: MIT Press.

- [105] Tiebout, C.M. (1956). "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64, 416-424.
- [106] Volden, C. (2005). "Intergovernmental Political Competition in American Federalism." *American Journal of Political Science*, 49, 327-342.
- [107] Volden, C. (2007). "Intergovernmental Grants: A Formal Model of Interrelated National and Subnational Political Decisions." *Publius: The Journal of Federalism*, 37, 209-243.
- [108] Weisstein, E. W. (1999). *The CRC Concise Encyclopedia of Mathematics*. CRC Press LLC.
- [109] Zupan, M.A. (1991). "An Economic Explanation for the Existence and Nature of Political Ticket Splitting." *Journal of Law and Economics*, 34, 343-369.